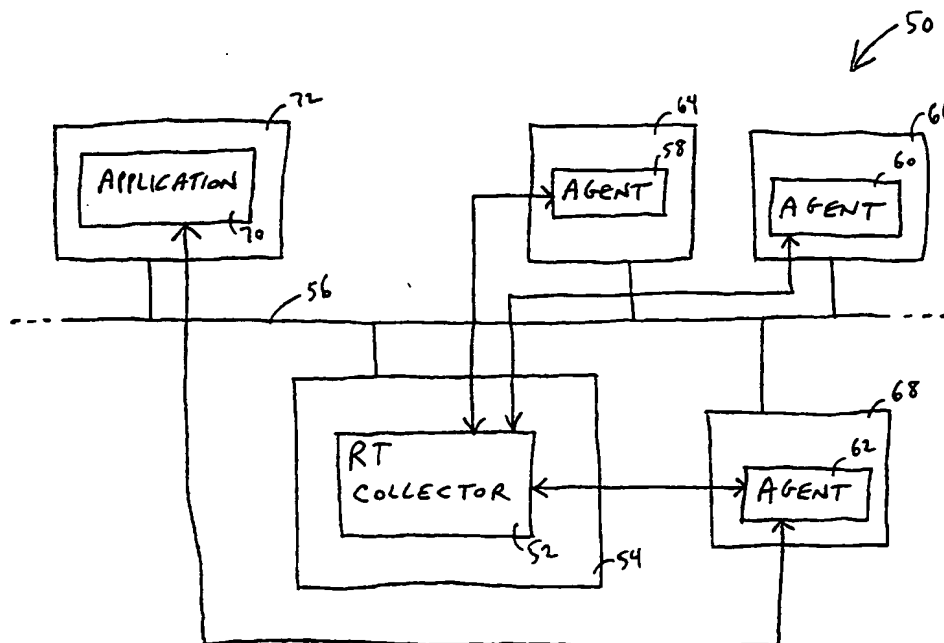




INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁷ : G06F 11/34, H04L 12/26		A1	(11) International Publication Number: WO 00/19320
			(43) International Publication Date: 6 April 2000 (06.04.00)
(21) International Application Number: PCT/US99/22787		(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).	
(22) International Filing Date: 30 September 1999 (30.09.99)		Published <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>	
(30) Priority Data: 60/102,579 30 September 1998 (30.09.98) US			
(71) Applicant: NEXTPOINT NETWORKS [US/US]; 68 Tadmuck Road, Westford, MA 01886 (US).			
(72) Inventors: SWEET, Bruce; 19 Kings Row, North Reading, MA 01864 (US). KELLEY, Bruce, A., Jr.; 8 Hyacinth Drive, Westford, MA 01886 (US). DARUWALLA, Gev; 242 Congress Street, Milford, MA 01757 (US). FULREADER, John; 160 Newport Street, Arlington, MA 02476 (US). PEGRAM, Gregory; 9 Dow Drive, Haverhill, MA 01832 (US).			
(74) Agents: REYES, Jason, A. et al.; Hale and Dorr LLP, 60 State Street, Boston, MA 02109 (US).			

(54) Title: EVALUATING COMPUTER RESOURCES BY END-USER EMULATION



(57) Abstract

To closely simulate the experience of an end user, a software agent executes where the end user would be situated and attempts to exercise a computer resource such as a networked application or a network resource in a simulated transaction in exactly or nearly exactly the same way that the computer resource would be exercised by the end user in non-simulated transaction. The results of the simulated transaction and of other simulated transactions by other software agents are communicated to a central software system for recordation and analysis.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

-1-

EVALUATING COMPUTER RESOURCES BY END-USER EMULATIONCross-Reference to Related Applications

This application claims the benefit of United States Provisional
5 Application Serial No. 60/102579 entitled "MANAGING COMPUTER
NETWORK RESOURCES" filed on September 30, 1998, which is incorporated
herein.

Background

10 The present invention relates to a method and apparatus for evaluating
computer resources.

To efficiently manage and troubleshoot computer resources such as
software applications on a network, an administrator needs information about
the status of the computer resources and communication paths on the network.
15 To effectively monitor the resources and the network, the administrator needs
information regarding the availability and performance of networked
applications and the ability of the network to provide access to the networked
applications. Such detailed information is important for management of the
computer resources and the network so that the configuration of the resources
20 or the network can be modified or supplemented as needed. Preferably, these
modifications or supplements are put in place before major problems occur.

-2-

Currently, the information that is available about a software application running on a server may be derived from information that is accessible to the server or the application, and the information that is available about a network may be derived from information gathered by a network analyzing product
5 such as a network sniffer, which is a device that is inserted into a segment of the network to monitor the packets that are traveling along the communication line into which the sniffer has been inserted.

The increasing popularity and complexity of Web site and networked database applications that are dependent on the Internet, or on a network
10 using Internet protocols, has further complicated the administrative task, including the task of managing Web site database content and maintaining Web site effectiveness. Company Webmasters and business managers are routinely faced with a wide array of burdensome issues such as the detection, identification, and repair of broken Web sites and links, the monitoring and
15 organization of large volumes of diverse, continuously-changing Web site and networked database content, and the detection and management of Web site problems and congested links. These issues are particularly significant for companies that rely on their Web sites to provide mission-critical information and services to customers and business partners.

-3-

Summary of the Invention

A method and apparatus are provided for evaluating computer resources. To closely simulate the experience of an end user, a software agent executes in a location where the end user would be situated and attempts to exercise a computer resource such as a networked application or a network resource in a simulated transaction in exactly or nearly exactly the same way that the computer resource would be exercised by the end user in a non-simulated transaction. The results of the simulated transaction and of other simulated transactions by other software agents are communicated to a central software system for recordation and analysis.

The method and apparatus may be used with the World-Wide Web or the Internet or with a local area network ("LAN") or a wide area network ("WAN") that extends over a large enterprise and includes a number of LANs connected together in an "intranet."

Different aspects of the invention allow one or more of the following. An end-to-end evaluation can be performed of the ability of an overall system, that includes an application running on a server and a network providing access to the application, to serve an end-user effectively. A pro-active detection system can be provided that detects a problem in the overall system before the problem affects, or is detected by, an end-user. Simulated transactions that are highly representative of non-simulated transactions can be executed using the same computer resources that are used by the non-

-4-

simulated transactions. Existing computer and network hardware already installed in a network can be used highly effectively to evaluate computer resources on the network. Evaluations of computer resources on a network can be performed at effectively any time with little or no disruption to the normal operation of the computer resources or the network.

Other features and advantages will become apparent from the following description, including the drawings, and from the claims.

Brief Description of the Drawings

Figs. 1-2 and 7 are block diagrams of computer-based systems.
Figs. 3A-6 are flow diagrams of computer-based procedures.

Detailed Description

A Synthetic Transactions system that simulates the actions of an actual distributed network software application may be used to evaluate computer resource (e.g., application) response time or availability over a computer network. In at least some cases, the evaluations are taken from the perspective of an end user (which may be a human being or a software program), so that the evaluations reflect the application response time and availability that would likely be experienced by the end user.

As described in detail below, a Synthetic Transaction is a simulated application transaction. In at least some cases, it is advantageous if the

-5-

simulated transaction is generated by an intelligent agent software system to evaluate a computer resource. For example, in a typical case concerning client/server technology on a network, a Synthetic Transactions software agent emulates a client side of a client/server transaction so that what happens on the network and on a server is the same, or nearly the same, as what would happen if the transaction were initiated by an actual end user on the client side. The agent may execute at a point on the network that is highly representative of where an actual end user would be situated, so that the path of associated communication across the network closely simulates the path of communication that is, or likely would be, associated with an actual end user.

Fig. 1 illustrates relationships among components of an example implementation ("NextPoint S3") of the Synthetic Transactions system, described in more detail below.

In the Synthetic Transactions system, an application response time is an amount of time between two events in the course of a transaction. Typically, the application response time is an amount of time between a time when a transaction involving an application is started by the software agent and a later time when the transaction is completed (e.g., when the last protocol message has been processed). However, intermediate points in the course of the transaction could be used where appropriate. Typically, the application response time includes any time spent in the transaction due to latencies in the computer network, the application server, any additional backend servers such

-6-

as a database server, as well as time spent by the software agent (e.g., in processing application protocol messages).

An overall application response time measured by a software agent may include multiple components, such as time spent on the network and time spent in the application, which may also include multiple components such as time spent for individual queries in a database application.

One or more of many methods could be used in determining how much of an overall application response time is attributable to network travel time ("network response time") and to time spent in processing by the application ("processing response time"), respectively. In a first example method, an Internet Control Message Protocol Packet Internet Groper ("ICMP Ping") protocol transmission is used to determine the network response time. An ICMP Ping protocol transaction transmission is sent simultaneously, or nearly simultaneously, with an application protocol message to the application server. At the application server, the ICMP Ping protocol transaction transmission is processed at a low level without involving the application software, to issue a response. The amount of time that passes between the sending of the ICMP Ping protocol transmission and the receiving of the response represents the network response time. The processing response time may be determined by subtracting the network response time from the overall application response time.

-7-

In a second example method, a Transmission Control Protocol/Internet Protocol ("TCP/IP") socket connection time is determined and used to determine the network response time. An initial step in executing any TCP/IP network application transaction is the establishment of a TCP/IP socket connection between the client and the server, which establishment is used in the second method to measure the network response time. The establishment of the socket connection does involve the application on the server, because the connection is not successful unless the application accepts the connection, but such involvement does not include hard disk access or back end server components and therefore takes up relatively insignificant time. The network response time is taken to be the TCP/IP socket connection time (i.e., the amount of time needed to establish the socket connection) and the processing response time is determined by subtracting the network response time from the overall application response time.

An advantage of the second example method is that the measurement is made using the same network path that is used in typical non-evaluative transactions between the client and the server. In the case of the first example method, an ICMP Ping transmission may be filtered out by an intermediate network processor such as a firewall, or a network device providing Quality of Service ("QOS") services in place within the computer network may cause the ICMP Ping transmission to be given a priority that is different from the priority that a typical non-evaluative transmission would be given. Such a

-8-

difference in priority may affect the network path followed or the processing time for the ICMP Ping transaction at the server, which would skew the evaluation of the network response time.

In a specific implementation, multiple methods, such as both the first and second example methods described above, may be made available for selection and execution.

Application availability is determined based on whether the application responds to a Synthetic Transactions transmission, regardless of the response time. If a response to the Synthetic Transactions transmission is received before a designated expiration time has passed, the application is noted as available. Otherwise, the application is noted as unavailable.

Synthetic Transactions may be executed on a scheduled basis such as a periodic basis, or on user demand. For example, a Synthetic Transaction may repeatedly exercise a specific application function, such as reading a customer record from a database. As a result, a baseline of expectations for the latency of the transaction being simulated may be established, and deviations from the baseline may be interpreted as indicating abnormal functioning of the computer network or the application. In cases in which the network response time and the processing response time are determined separately and corresponding separate baselines are established, the abnormal functioning may be isolated to the network or the application, respectively. See the

-9-

commonly assigned, simultaneously filed United States Patent Application entitled "MANAGING COMPUTER RESOURCES".

Synthetic Transactions may be executed for at least the following network application technologies: ICMP Ping, Dynamic Host Control Protocol ("DHCP"), Domain Name Service ("DNS"), Post Office Protocol 3 ("POP3"), Simple Mail Transfer Protocol ("SMTP"), Standard Query Language ("SQL") Database queries, Hypertext Transport Protocol ("HTTP") web applications, eCommerce, SAP/R3, PeopleSoft, Lotus Notes, Microsoft Exchange and Voice over IP. Synthetic Transactions are not limited to Internet Protocol network applications and may be used for Systems Network Architecture ("SNA"), Internet Protocol Exchange ("IPX"), or Appletalk network applications, for example. All or a part of a network application may be evaluated through the use of Synthetic Transactions. In a typical implementation, a set of Synthetic Transactions is developed to evaluate common and important operations performed by the application. The user may be required to supply parameters that specify a particular instance of the application to be used. For example, the parameters may include the name of a server and a username and password to log into an application on the server.

In at least some cases, it may be advantageous if each set of Synthetic Transactions for an application has characteristics that are particularly appropriate for that application. Although the software agent may provide a common infrastructure for all transactions, the transactions for evaluations of a

-10-

particular application may be tied to that application. For example, SQL scripts are supported by Synthetic Transactions for database queries. The SQL scripts allow for effectively infinite flexibility in the number and type of database related Synthetic Transactions that may be executed. In another example, a
5 Synthetic Transaction to evaluate an ability to look up an Internet Domain Name by the Domain Name Service may have characteristics that render the transaction significantly less useful for another type of evaluation.

In at least some cases, the software agent generates a Synthetic Transaction by creating a set of network protocol messages that are
10 appropriate for an application that is being evaluated. For example, as described in more detail below, to measure response time concerning a Web application, the software agent may build and generate HTTP protocol messages, to emulate a Web browser function. In such a case, a specific Uniform Resource Locator ("URL") identifier for a Web page is encoded in one
15 or more of the HTTP protocol messages and the software agent attempts to retrieve all or a portion of the Web page specified by the URL identifier.

Information generated from a Synthetic Transaction may be temporarily stored by the software agent and may be subsequently sent to a central server for long term storage and for further processing and analysis. The software
20 agent may be controlled by software at the central server, for a purpose such as so that the software agent is specifically directed to generate information needed by the central server for analysis of the network or computer resources.

-11-

For example, software at the central server may determine a set of Synthetic Transactions to be executed and the schedule for the execution of such transactions, and then may apply the set and the schedule to the software agent, which may then execute autonomously to generate the information
5 needed by the software at the central server. (In certain implementations, the software agent may be able to create, specify, or vary the set or the schedule, perhaps in response to varying conditions or inputs.)

Synthetic Transactions may be used to measure throughput, e.g., by keeping track of how much time is required for an amount of data to pass
10 through a particular point in the network.

A specific example of a Synthetic Transaction is described below. Fig. 2 illustrates a network system 50 in which server software ("RT Collector") 52 running on a computer 54 communicates on a network 56 with software agents 58, 60, 62 ("Agents") running on other computers 64, 66, 68. The RT Collector
15 and Agents execute as described below to conduct transactions in connection with evaluating a networked application 70 running on a Web site 72.

The RT Collector is initialized (Figs. 3A-3B) (step 1010), and sets up a timer for the transactions that the RT Collector is to direct the Agents to execute (step 1020). The timer is set to issue a signal at periodic intervals
20 (which may be user defined) when transactions are to be executed.

When the timer signals (step 1030), the RT Collector determines whether the RT Collector is currently in contact with the Agent that is responsible for

-12-

performing a particular transaction (step 1040). If the RT Collector is not in contact with the Agent, the RT Collector opens a TCP socket connection to the Agent on a specified port such as port 9000 (step 1050). (If the RT Collector is unable to establish or re-establish a connection to the Agent, the RT Collector

5 tries to establish the connection every time that the timer signals that the Agent is to perform a transaction.) Once established, the connection between the Agent and the RT Collector is used for all communications between the RT Collector and the Agent until the connection is lost, which may occur due to a network failure or when the client or the server closes the connection.

10 The RT Collector and the Agent determine each other's version numbers as follows (step 1060), so that the RT Collector and the Agent are aware of each other's capabilities. The RT Collector submits, to the Agent, a version request packet that indicates the version number of the RT Collector, and waits a fixed amount of time for the Agent to return another version packet that

15 indicates the version number of the Agent. If the other version packet is not received before the fixed amount of time passes, the RT Collector assumes that the Agent has a version number, e.g., "1.0.0.0.", that corresponds to a basic set of capabilities.

After the version numbers are determined, the RT Collector determines

20 all of the timers that have signaled (step 1070) and identifies a set of transactions corresponding to the timers (step 1080). The RT Collector determines whether a previous instance of any transaction in the set is

-13-

currently being performed by the Agent (step 1090). If so, the RT Collector determines that the previous instance was not performed successfully (which may occur for a reason such as a malfunction in the Agent) (step 1100) and generates an exception indicating that the Agent is unavailable (step 1110).

- 5 Otherwise, the RT Collector submits information for the set of transactions to the Agent (step 1120).

The information for the set of transactions to be performed is sent to the Agent in the form of packets. Each packet has a common header that includes the following information. A transaction ID is included that is to be returned
10 to the RT Collector in response packets from the Agent so that the RT Collector can identify the transaction with which the response is associated. A transaction type ID is included to identify the type of transaction that the Agent is to perform (e.g., Web site transaction, e-commerce transaction, database query transaction). An integer is included to indicate the number of
15 times that the transaction is to be performed.

The remainder of each packet is specific to the transaction type. For example, in the case of a Web site transaction, the remainder may include the URL of a Web page at the Web site and an integer indicating the number of lines to read from the Web page.

- 20 Each response packet sent from the Agent to the RT Collector includes a transaction ID, which in a specific implementation is the first piece of data in the response packet. The remainder of the response packet may be specific to

-14-

the transaction type, and may include a response time value indicating a number of milliseconds of response time. A negative response time value may be used to indicate that an error occurred during performance of the transaction, and may include an error code bit mask that specifies the nature of the error to the RT Collector.

5 The Agent executes as follows (Figs. 4A-4B). The Agent interprets, and executes according to, any customized environment variables that may have been specified, such as variables included on a command line command or in a parameter file maintained on a hosting computer system on which the Agent executes (step 2010).

The Agent waits for a connection attempt on port 9000 (step 2020). When an RT Collector connects, a "Com Thread" thread of execution is established to handle communication between the Agent and the RT Collector (step 2030) and the Agent waits for another connection attempt on port 9000.

15 The Com Thread executes in a loop that includes reading a transaction request packet received from the RT Collector (step 2040), determining the type of transaction that is indicated in the request packet (step 2050), and establishing a Transaction Thread to perform the transaction (step 2060). The Transaction Thread receives, as parameters, a reference to the transaction request packet and a reference to the socket connection, which allows the Transaction Thread to respond directly to the RT Collector independently of other Agent behavior.

20

-15-

In a specific implementation, the Transaction Thread executes as follows. Transaction-specific information is extracted from the transaction request packet (step 2070), and the transaction is performed to produce results (step 2080). A response packet is produced that includes the results (step 5 2090), and the response packet is sent to the RT Collector (step 2100).

A Web site transaction may be performed as follows (Figs. 5A-5D), in a specific implementation. The URL of a Web site page to be read and a number specifying the number of lines of the page to be read are extracted from the transaction request packet (step 3010). The current time is recorded as a transaction start time (step 3020). An instance of a Java URL class is 10 created based on the URL (step 3030). A network connection is created for the URL (step 3035), and options are set on the network connection, such as option concerning a selectable address or a customization in connection with the computing environment (step 3040). If the user has requested that a 15 password protected HTTP proxy server be used, the encrypted password is specified (step 3050). The URL is designated to be for a Web page that is not to be fetched from any cache (step 3060).

An input stream is opened on the network connection (step 3070), and an input stream reader is opened on the input stream (step 3080). A watchdog 20 timer is created that causes the evaluation to be halted if an expected response to any read request is not received within a specified amount of time, such as 30 seconds (step 3090).

-16-

The following steps 3110, 3120 are performed until the specified number of lines of the Web page are read, an error occurs, or all of the lines of the Web page have been read (step 3100). A line of the Web page is read (step 3110). To prevent the watchdog timer from expiring, the watchdog timer is
5 reset (step 3120).

Once the specified number of lines have been read, the watchdog timer is terminated (step 3130). The current time is recorded as a transaction end time (step 3140). All transaction specific resources are cleared (step 3150).

An elapsed response time is calculated by determining the difference
10 between the transaction end time and the transaction start time (step 3160). A response packet is built that includes the transaction ID that was passed in the request packet and either the elapsed response time or an error code as described above (step 3170). The response packet is sent to the RT Collector (step 3180) and the Transaction Thread is terminated (step 3190).

15 If the watchdog timer expires, the resources associated with the transaction are terminated (step 3200) and the RT Collector is sent a response packet indicating that the transaction timed out (step 3210).

The RT Collector executes as follows upon receiving a response packet from the Agent (Fig. 6). The response packet is appended to a queue of
20 response packets that is configured to hold the response packets of all of the Agents to which the RT Collector submits transaction requests (step 4010). In turn, the response packet is retrieved from the queue of response packets (step

-17-

4020). The transaction with which the response is associated is identified (step 4030) and is determined to be concluded (step 4040). If the results of the transaction indicate an error, an exception is posted (step 4050).

5 If the hour of the day (or other appropriate interval) has changed, the results of all of the previous hour's instances of the transaction are added to the database (step 4060).

The results of the transaction are stored together with results for other instances of the transaction in the same time period, such as in the current hour (step 4070). If the transaction is executed more often than a particular
10 frequency, such as every 15 minutes, the corresponding results may be averaged (step 4080).

The technique (i.e., the procedures described above) may be implemented in hardware or software, or a combination of both. In at least some cases, it is advantageous if the technique is implemented in computer
15 programs executing on one or more programmable computers, such as a personal computer running or able to run Microsoft Windows 95, 98, or NT, Sun Solaris, LINUX, Unix, or MacIntosh OS, that each include a processor, a storage medium readable by the processor (including volatile and non-volatile memory and/or storage elements), at least one input device such as a
20 keyboard, and at least one output device. Program code is applied to data entered using the input device to perform the method described above and to

-18-

generate output information. The output information is applied to one or more output devices such as a display screen of the computer.

In at least some cases, it is advantageous if each program is implemented in a high level procedural or object-oriented programming language such as C++ or Java to communicate with a computer system. However, the programs can be implemented in assembly or machine language, if desired. In any case, the language may be a compiled or interpreted language.

In at least some cases, it is advantageous if each such computer program is stored on a storage medium or device, such as ROM or magnetic diskette, that is readable by a general or special purpose programmable computer for configuring and operating the computer when the storage medium or device is read by the computer to perform the procedures described in this document. The system may also be considered to be implemented as a computer-readable storage medium, configured with a computer program, where the storage medium so configured causes a computer to operate in a specific and predefined manner.

Other embodiments are within the scope of the following claims. For example, transactions may also be performed for e-commerce applications or database queries or for evaluations of network throughput or mass storage performance. In the case of database queries, a database transaction system

-19-

can be arranged in which SQL statements are submitted by a software agent to a remote database server on a periodic basis such as every five minutes.

The software agent may execute tasks such as breaking down the test results into components including the connection time, and automatically
5 executing a transaction multiple times at predetermined intervals in response to a single directive from the RT Collector. Execution threads may be arranged so that multiple RT Collectors may be served by a single software agent, and a single RT Collector may be served by multiple software agents. The software agent may acknowledge receipt of a request packet from the RT
10 Collector by sending an acknowledgement packet to the RT Collector. The software agent may have a class structure, such as the structure illustrated in Fig. 7, within which processing tasks are delegated. The class structure may include a facility for processing the results of a transaction, and perhaps also a facility that generates an alarm, an alert, or an exception when the processing
15 indicates that the results merit drawing attention to a computer resource involved in the transaction.

Transactions may be grouped by transaction IDs. Java based classes of distributed processes may be used, which may provide significant platform independence.

20 The software agent may maintain a data storage cache in which to store the results of multiple transactions, which may reduce the amount or frequency of communication between the software agent and the RT Collector

-20-

and may allow the software agent to operate autonomously during a period in which the RT Collector cannot communicate with the software agent. If the cache becomes full or nearly full, information in the cache may be discarded on a first in, first out basis.

5 The RT Collector may serve as a socket client. Java archive objects, SNMP Management Information Base ("MIB"), Extensible Markup Language ("XML") page description technology, or Push technology may be used, and may allow the software agent to supply results to the RT Collector or to another destination on a Push technology basis.

10 The transaction specific information supplied by the RT Collector may be derived from information that the RT Collector extracts from a database of evaluation transaction setup information.

-21-

What Is Claimed Is:

1. A method comprising:
causing a transaction to be executed that emulates an end-user's
interaction with a software application, wherein the transaction includes
5 communication with the software application across a network.
2. The method of claim 1, further comprising:
causing a message to be transmitted, to a point where the software
application is executing, from an end-user point on the network.
10
3. The method of claim 1, further comprising:
causing a message to be transmitted, from a point where the software
application is executing, to an end-user point on the network.
- 15 4. The method of claim 1, wherein the software application includes a
Web server application.
5. The method of claim 4, wherein:
causing the transaction to be executed includes emulating a function of
20 a Web browser.

-22-

6. The method of claim 1, wherein the network includes client/server technology and the transaction includes a client/server transaction.

7. The method of claim 6, further comprising:
emulating the client side of the client/server transaction.

8. The method of claim 1, further comprising
attempting to retrieve specified information from the software application.

9. The method of claim 1, further comprising
attempting to retrieve information from a Web page supplied by the software application.

10. The method of claim 1, further comprising:
supplying login information to the software application.

11. The method of claim 1, wherein the software application includes a database application.

12. The method of claim 11, further comprising:
attempting to retrieve information from the database application.

-23-

13. The method of claim 1, further comprising:
determining a response time in the transaction.

5 14. The method of claim 13, further comprising:
determining an application processing time component of the response
time.

10 15. The method of claim 13, further comprising:
determining a network processing time component of the response time.

15 16. The method of claim 1, further comprising:
attempting to establish a TCP/IP connection; and
based on a result of the attempt, determining a network processing
time.

17. The method of claim 1, further comprising:
transmitting a ping message; and
based on a result of the transmission, determining a network processing
time.

20 18. The method of claim 1, wherein the emulating includes emulating a
TCP/IP client.

-24-

19. The method of claim 1, further comprising:
based on the transaction, determining the availability of the software application.

5 20. The method of claim 19, further comprising:
if a response is not received timely from the software application,
determining that the software application is unavailable.

21. The method of claim 1, further comprising:
10 causing multiple instances of the transaction to be executed according to
a schedule.

22. The method of claim 21, further comprising:
deriving a performance baseline from results of the multiple instances.

15 23. The method of claim 1, further comprising:
causing, from a central location, multiple emulation transactions to be
executed.

20 24. The method of claim 23, wherein at least two of the transactions
originate from two different respective locations on the network.

-25-

25. The method of claim 24, further comprising:

storing information derived from a result of at least one of the transactions at the respective originating location for later transmission to the central location.

5

26. The method of claim 1, further comprising:

transmitting a data packet to the software application, the data packet including information that identifies the transaction.

10

27. A method comprising:

generating a simulated computer communication at a point in a computer network;

using a non-simulated network resource to deliver the simulated computer communication to a computer resource; and

15

evaluating the computer resource based on a result of the delivery of the simulated computer communication to the computer resource.

20

28. The method of claim 27, wherein the point in the computer network includes an end-user point.

29. The method of claim 27, further comprising:

-26-

issuing, from a central source, a directive to generate the simulated computer communication; and

receiving, at the central source, a report on the result of the delivery of the simulated computer communication to the computer resource.

5

30. Computer software, residing on a computer-readable storage medium, comprising a set of instructions for use in a computer system to cause the computer system to:

cause a transaction to be executed that emulates an end-user's
10 interaction with a software application, wherein the transaction includes communication with the software application across a network.

31. Computer software, residing on a computer-readable storage medium, comprising a set of instructions for use in a computer system to
15 cause the computer system to:

generate a simulated computer communication at a point in a computer network;

use a non-simulated network resource to deliver the simulated computer communication to a computer resource; and
20 evaluate the computer resource based on a result of the delivery of the simulated computer communication to the computer resource.

-27-

32. A system comprising:

a system that causes a transaction to be executed that emulates an end-user's interaction with a software application, wherein the transaction includes communication with the software application across a network.

5

33. A system comprising:

a generator that generates a simulated computer communication at a point in a computer network;

a utilizer that uses a non-simulated network resource to deliver the

10 simulated computer communication to a computer resource; and

an evaluator that evaluates the computer resource based on a result of the delivery of the simulated computer communication to the computer resource.

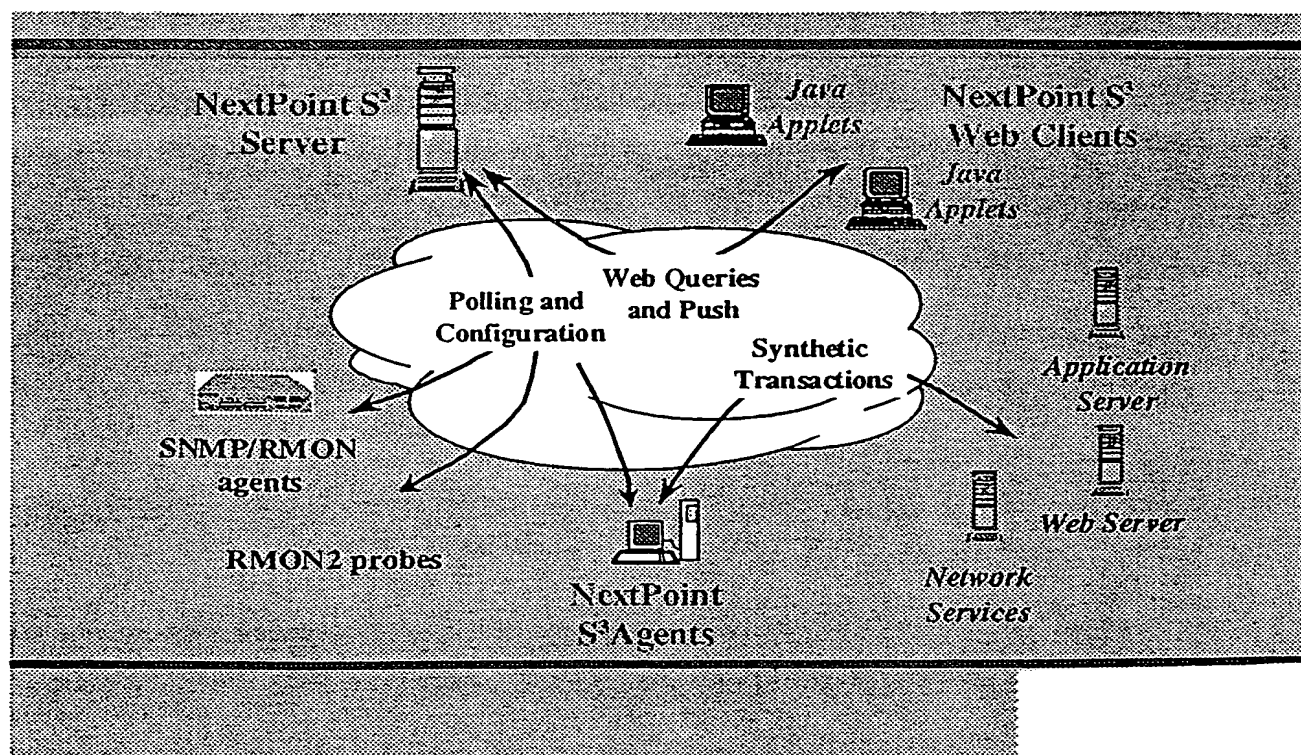
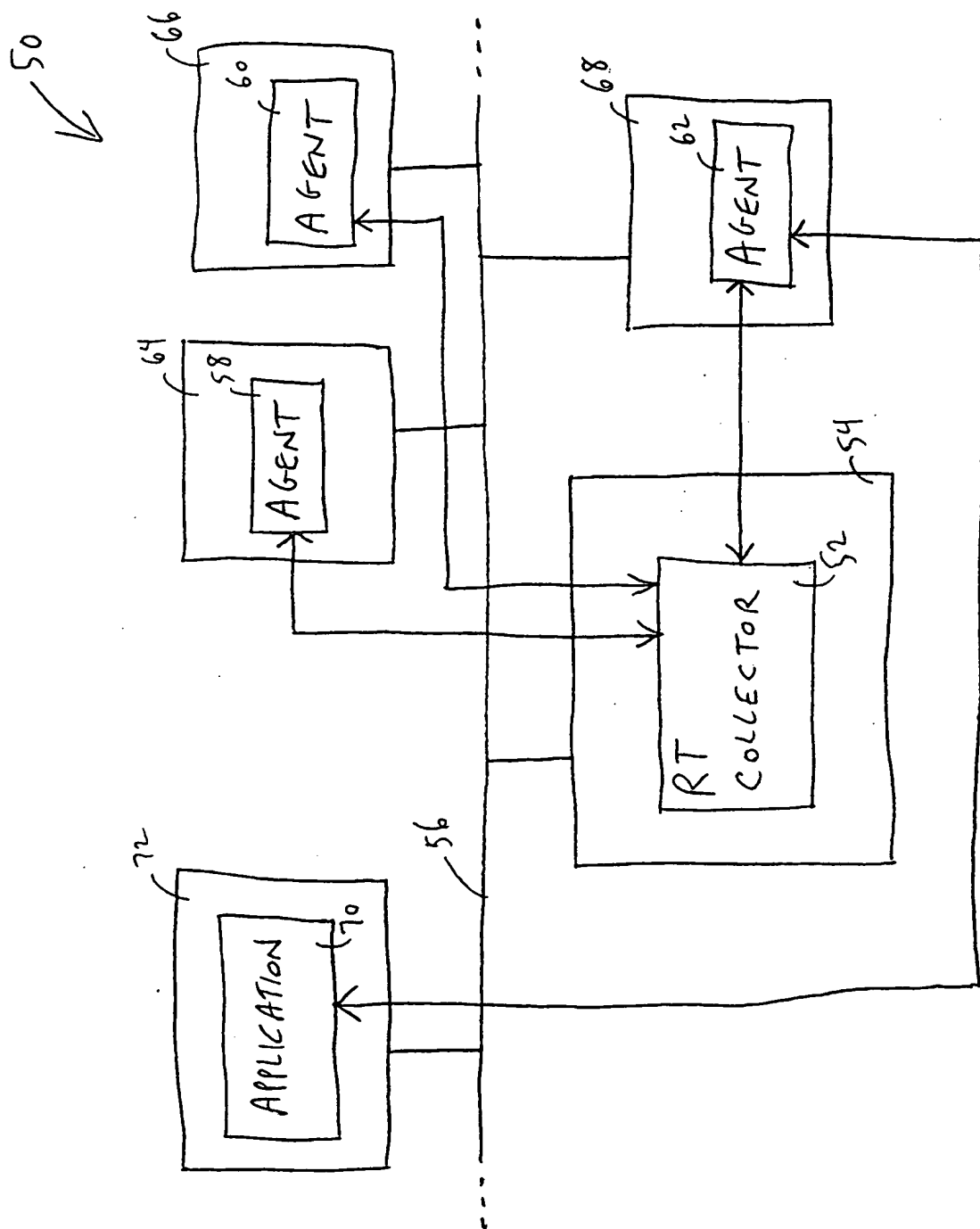


FIG. 1



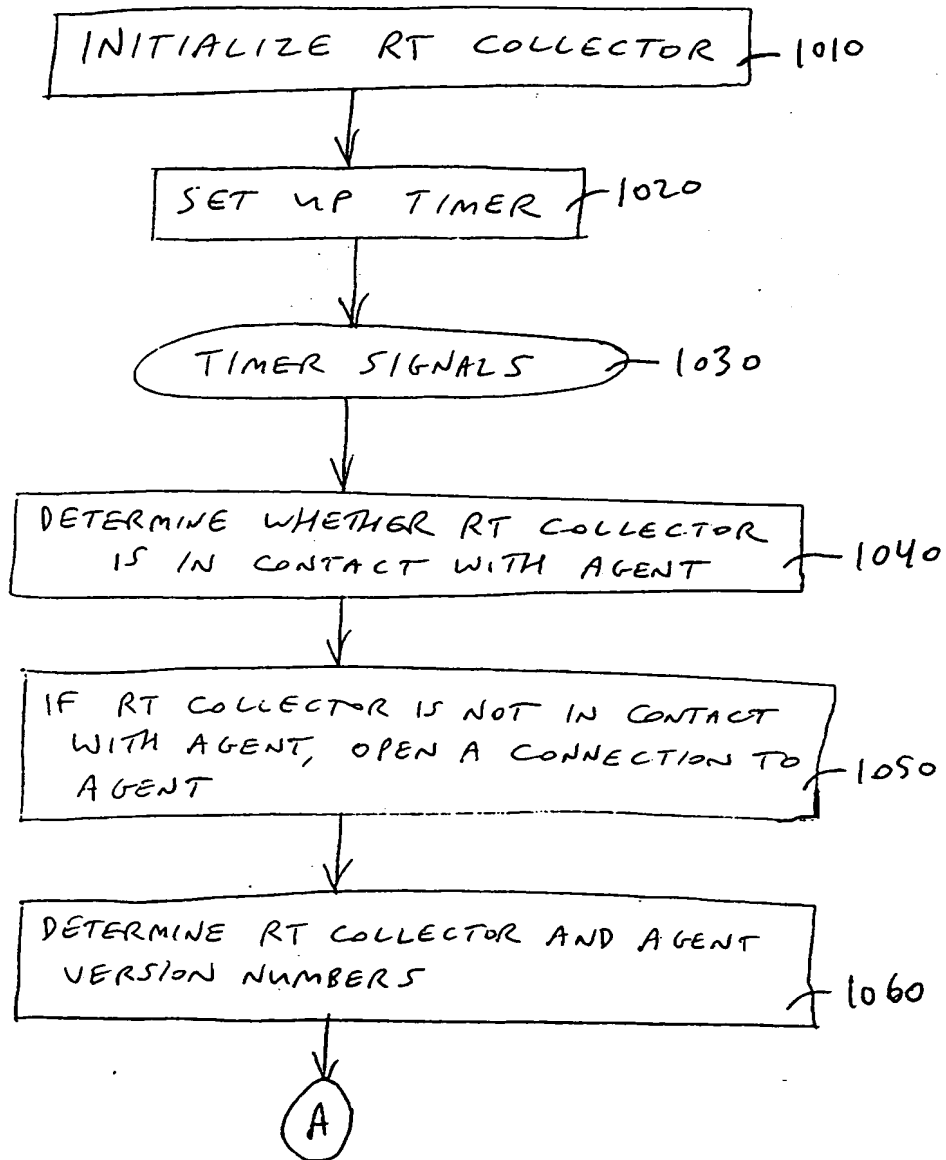


FIG. 3A

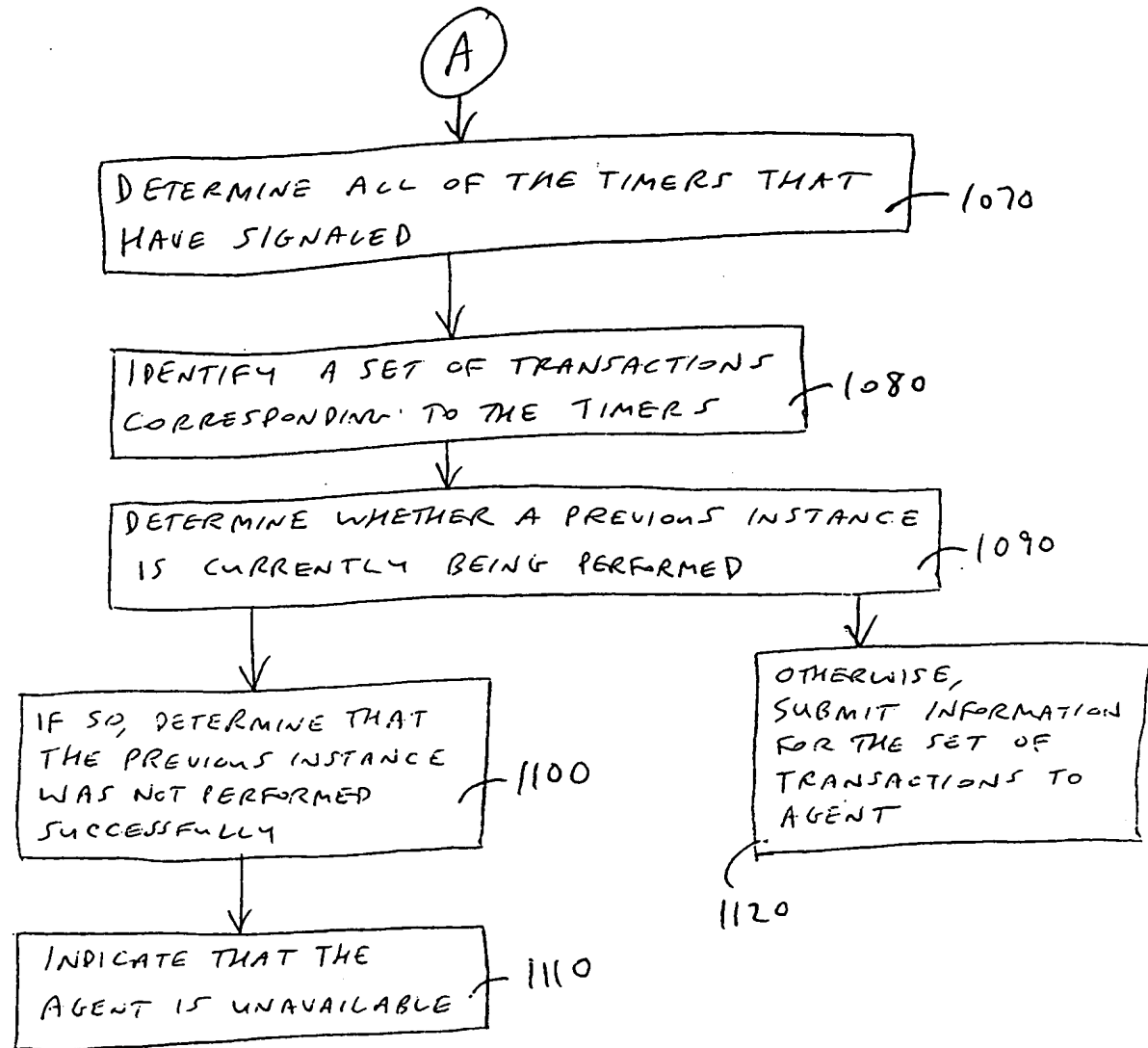


FIG. 3B

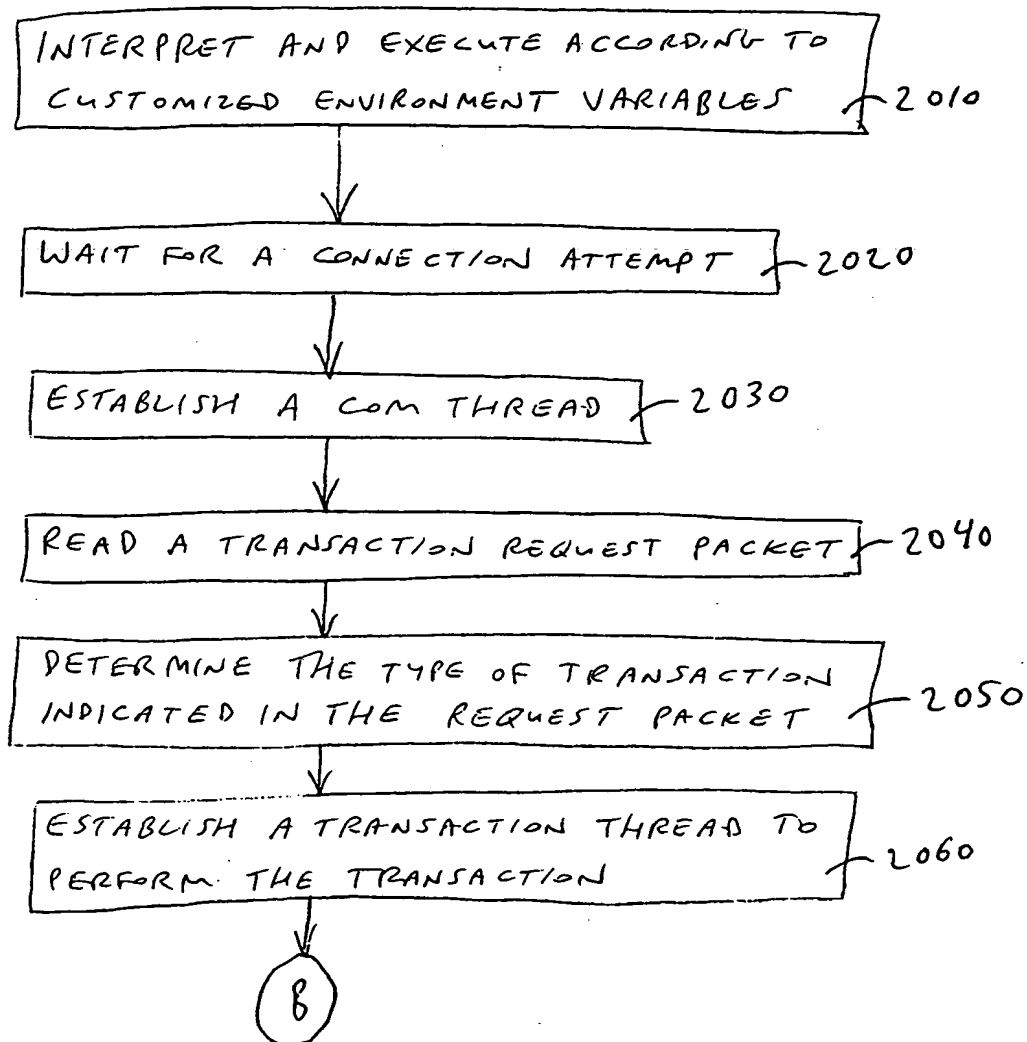


FIG. 4A

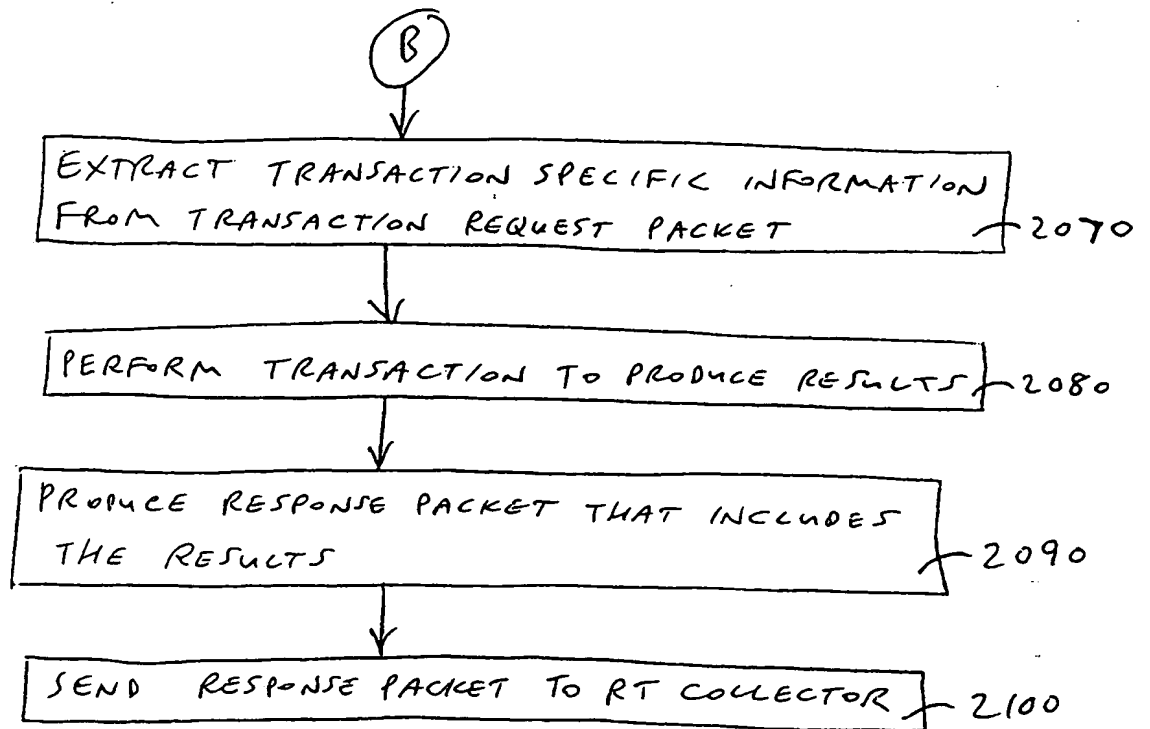


FIG. 4B

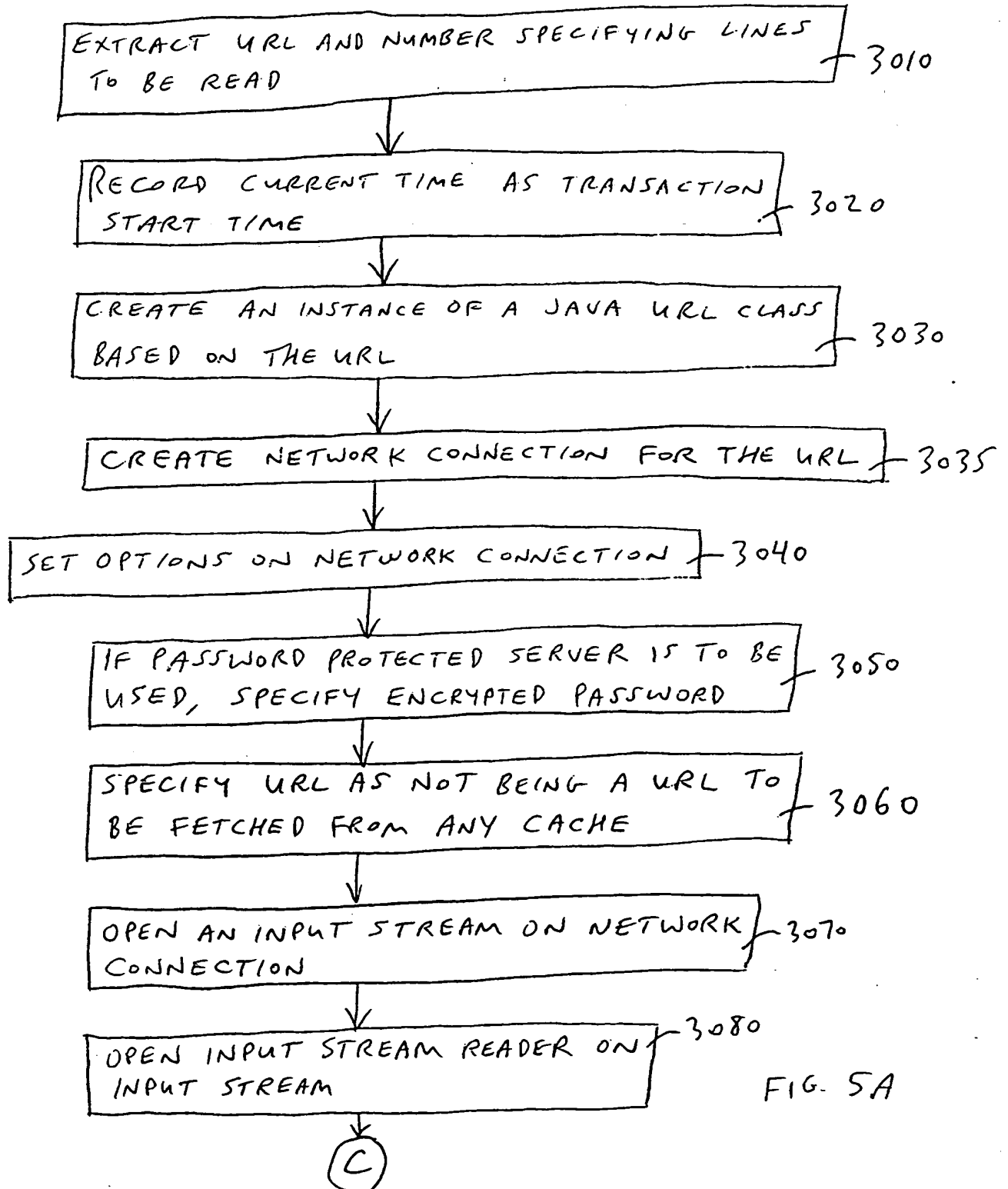


FIG. 5A

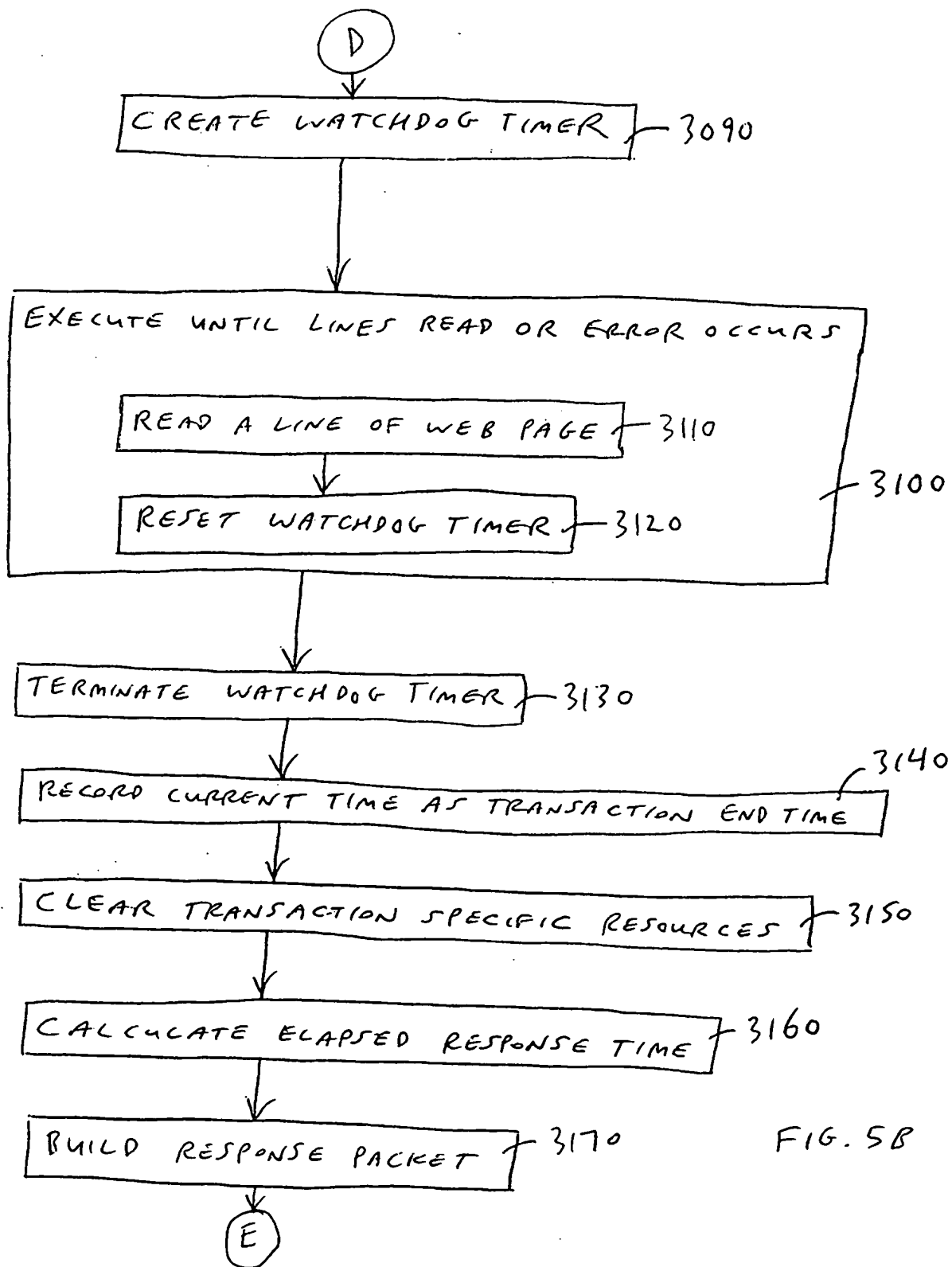


FIG. 5B

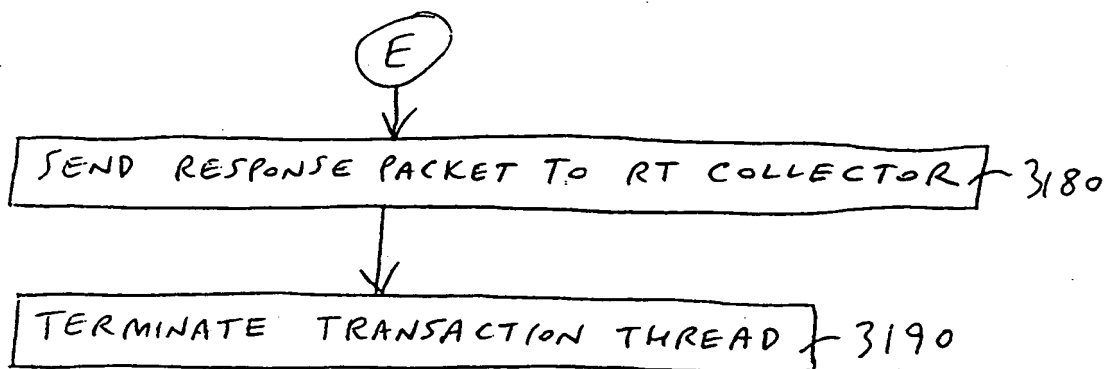


FIG. 5C

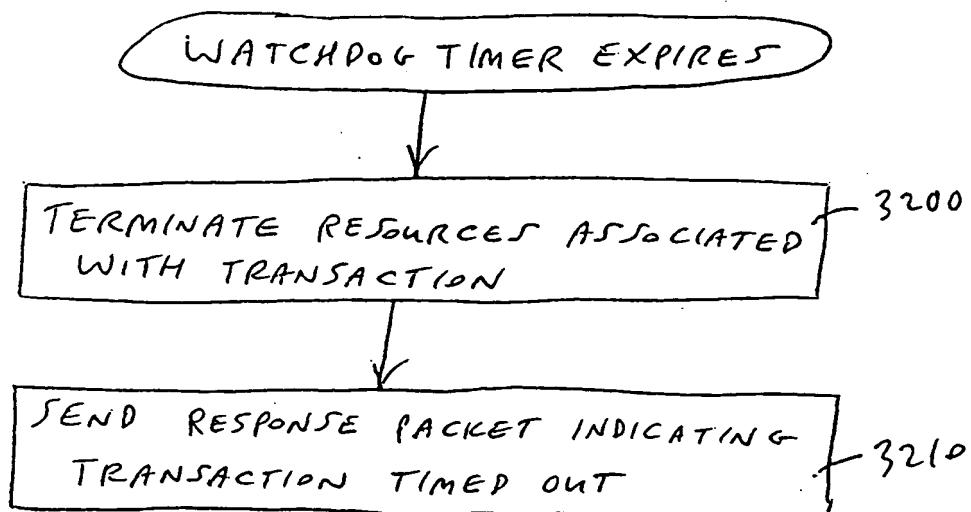
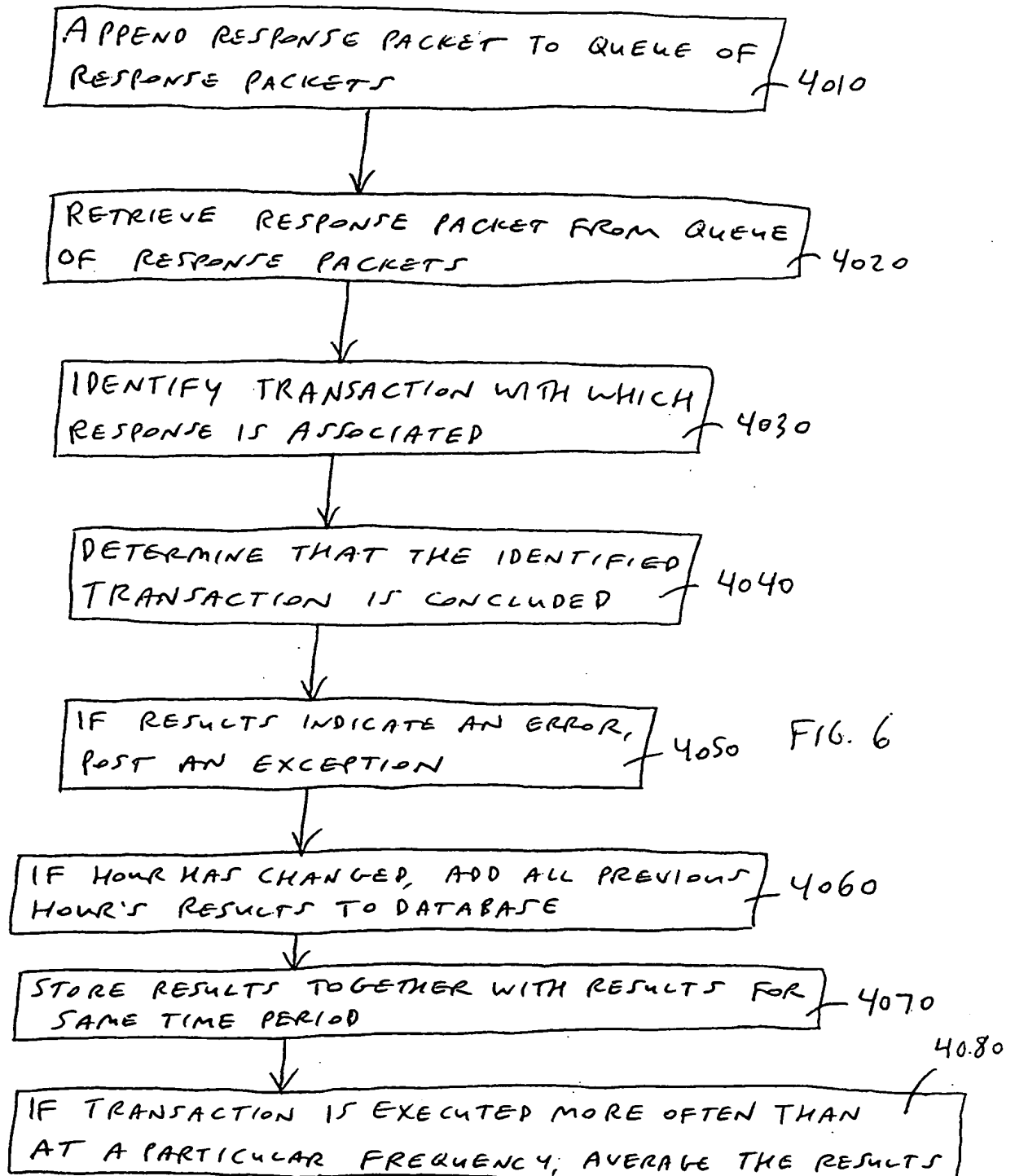


FIG. 5D



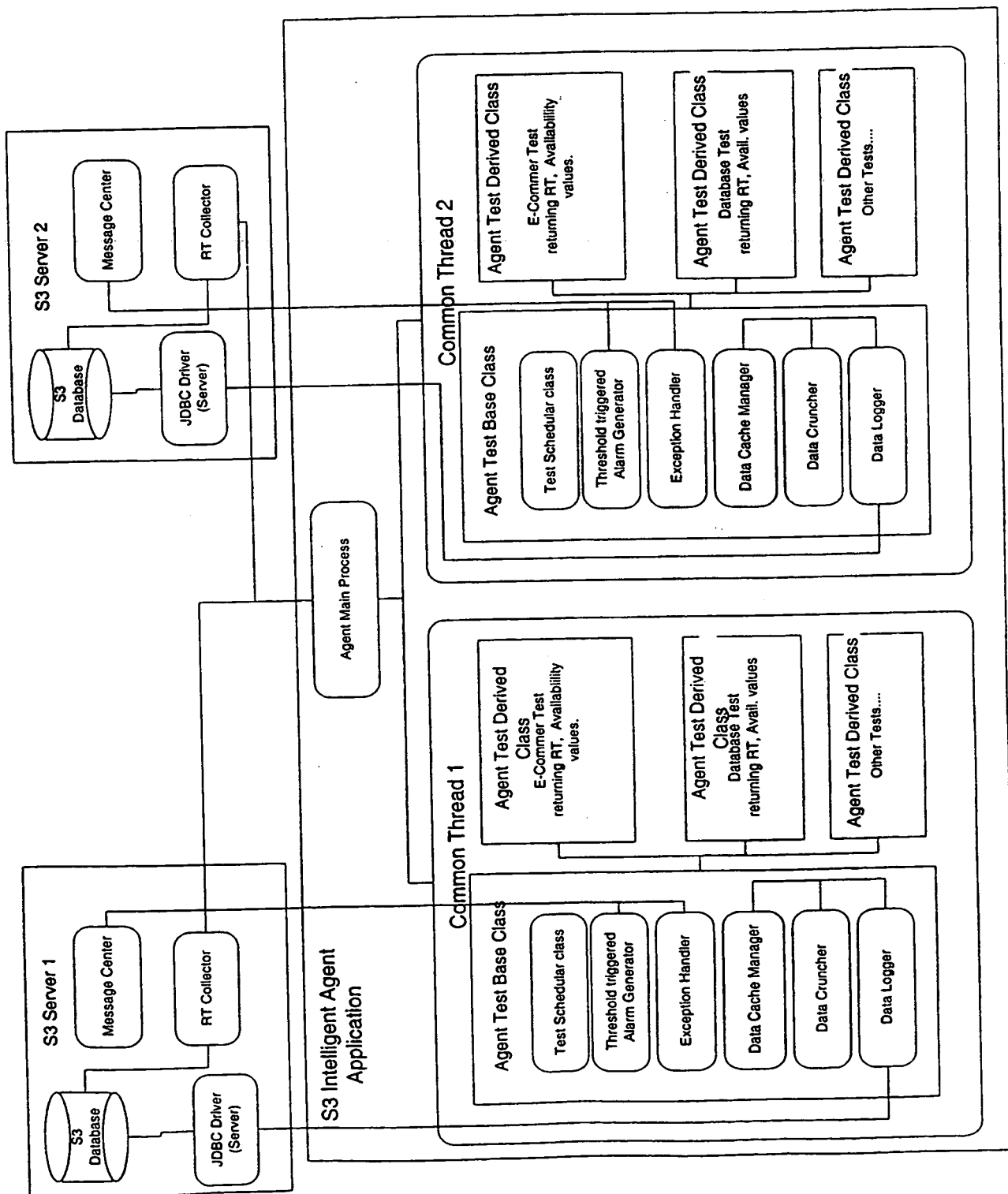


FIG. 7

INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 99/22787

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 G06F11/34 H04L12/26

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 G06F H04L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5 812 780 A (CHEN JOHN YUN-KUANG ET AL) 22 September 1998 (1998-09-22) column 1, line 1 -column 9, line 14 column 13, line 45 -column 14, line 38	1-4, 6-8, 10-13, 18-24, 26-33
X	US 5 577 197 A (BECK LONNIE P) 19 November 1996 (1996-11-19) the whole document	1-3, 6-8, 11-15, 19-21, 23-33
X	GB 2 307 318 A (MITSUBISHI ELECTRIC CORP) 21 May 1997 (1997-05-21) claim 1	1-3, 6-8, 11-13, 21-33
	— — — — — -/-	

☒ Further documents are listed in the continuation of box C.☒ Patent family members are listed in annex.

* Special categories of cited documents:

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
- "&" document member of the same patent family

Date of the actual completion of the international search

21 February 2000

Date of mailing of the international search report

06/03/2000

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3018

Authorized officer

Herreman, G

INTERNATIONAL SEARCH REPORT

International Application No.

PCT/US 99/22787

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	PATENT ABSTRACTS OF JAPAN vol. 006, no. 137 (P-130), 24 July 1982 (1982-07-24) & JP 57 059257 A (AGENCY OF IND SCIENCE & TECHNOL), 9 April 1982 (1982-04-09) abstract	1-3,6,7, 13-15, 21,23, 24,27-33
A	BORELLA M S ET AL: "WANDS: Wide-Area Network Delay Simulator" PROCEEDING OF THE IEEE 7TH INTERNATIONAL WORKSHOP ON NETWORK AND OPERATING SYSTEMS SUPPORT FOR DIGITAL AUDIO AND VIDEO (CAT. NO.97TH8259), PROCEEDINGS OF 7TH INTERNATIONAL WORKSHOP ON NETWORK AND OPERATING SYSTEM SUPPORT FOR DIGITAL AUDIO AND VIDEO, pages 85-92, XP002131053 1997, New York, NY, USA, IEEE, USA ISBN: 0-7803-3799-9 page 87, paragraph 3.1CRAWLER	1,4,5,9
A	LARSEN A K: "ALL EYES ON IP TRAFFIC. NEW APPS CAN MONITOR INTERNET AND INTRANET TRAFFIC, BUT DO THEY DELIVER ENOUGH DATA TO HOLD ISPS TO THEIR PROMISES?" DATA COMMUNICATIONS,US,MCGRW HILL. NEW YORK, vol. 26, no. 4, 21 March 1997 (1997-03-21), pages 54,56-60,62, XP000659549 ISSN: 0363-6399 page 59, middle column, line 13 -page 60, left-hand column, line 14	16,17

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 99/22787

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 5812780 A	22-09-1998	NONE	
US 5577197 A	19-11-1996	NONE	
GB 2307318 A	21-05-1997	JP 9138777 A US 5889955 A	27-05-1997 30-03-1999
JP 57059257 A	09-04-1982	JP 1323657 C JP 60053891 B	27-06-1986 27-11-1985

CORRECTED VERSION

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
6 April 2000 (06.04.2000)

PCT

(10) International Publication Number
WO 00/19320 A1

(51) International Patent Classification⁷: G06F 11/34,
H04L 12/26

(21) International Application Number: PCT/US99/22787

(22) International Filing Date:
30 September 1999 (30.09.1999)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/102,579 30 September 1998 (30.09.1998) US

(71) Applicant: NETSCOUT SERVICE LEVEL CORPO-
RATION [US/US]; 4 Technology Park Drive, Westford,
MA 01886 (US).

(72) Inventors: SWEET, Bruce; 19 Kings Row, North Read-
ing, MA 01864 (US). KELLEY, Bruce, A., Jr.; 8 Hyacinth

Drive, Westford, MA 01886 (US). DARUWALLA, Gev;
242 Congress Street, Milford, MA 01757 (US). FUL-
READER, John; 160 Newport Street, Arlington, MA
02476 (US). PEGRAM, Gregory; 9 Dow Drive, Haver-
hill, MA 01832 (US).

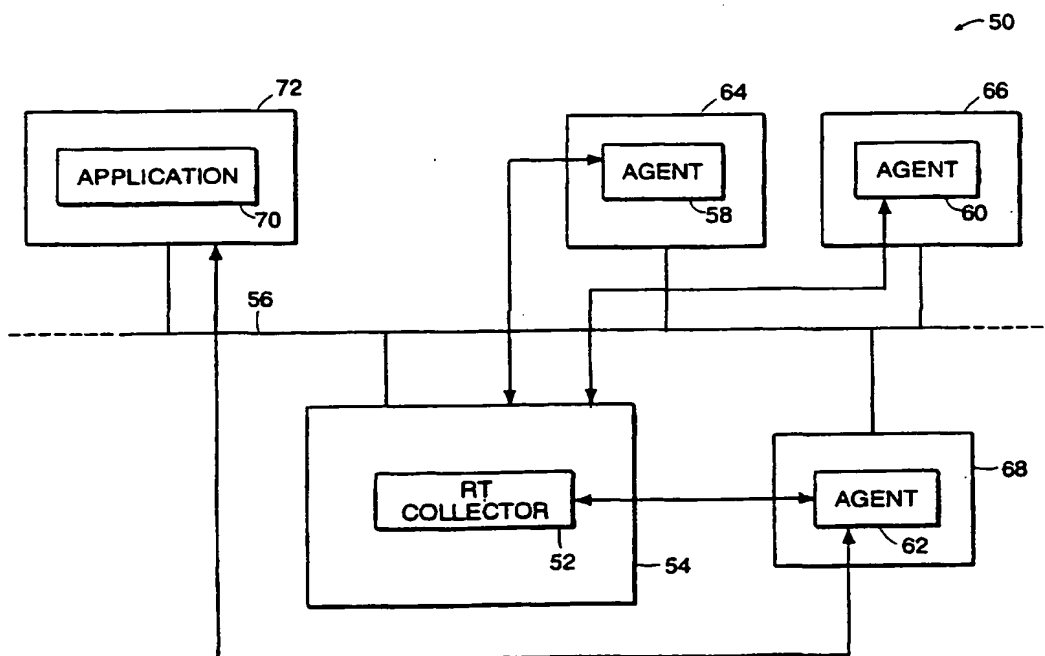
(74) Agent: FRANK, Steven, J.; Testa, Hurwitz & Thibault,
LLP, High Street Tower, 125 High Street, Boston, MA
02110 (US).

(81) Designated States (*national*): AL, AM, AT, AU, AZ, BA,
BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES,
FI, GB, GE, GH, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC,
LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX,
NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM,
TR, TT, UA, UG, UZ, VN, YU, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM,
KE, LS, MW, SD, SL, SZ, TZ, UG, ZW). Eurasian patent
(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent
(AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU,

[Continued on next page]

(54) Title: EVALUATING COMPUTER RESOURCES BY END-USER EMULATION



(57) Abstract: To closely simulate the experience of an end user, a software agent executes where the end user would be situated and attempts to exercise a computer resource such as a networked application or a network resource in a simulated transaction in exactly or nearly exactly the same way that the computer resource would be exercised by the end user in non-simulated transaction. The results of the simulated transaction and of other simulated transactions by other software agents are communicated to a central software system for recordation and analysis.



WO 00/19320 A1



MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

(15) Information about Correction:

see PCT Gazette No. 17/2002 of 25 April 2002, Section II

Published:

— with international search report

(48) Date of publication of this corrected version:

25 April 2002

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

-1-

EVALUATING COMPUTER RESOURCES BY END-USER EMULATIONCross-Reference to Related Applications

This application claims the benefit of United States Provisional
5 Application Serial No. 60/102579 entitled "MANAGING COMPUTER
NETWORK RESOURCES" filed on September 30, 1998, which is incorporated
herein.

Background

10 The present invention relates to a method and apparatus for evaluating
computer resources.

To efficiently manage and troubleshoot computer resources such as
software applications on a network, an administrator needs information about
the status of the computer resources and communication paths on the network.
15 To effectively monitor the resources and the network, the administrator needs
information regarding the availability and performance of networked
applications and the ability of the network to provide access to the networked
applications. Such detailed information is important for management of the
computer resources and the network so that the configuration of the resources
20 or the network can be modified or supplemented as needed. Preferably, these
modifications or supplements are put in place before major problems occur.

-2-

Currently, the information that is available about a software application running on a server may be derived from information that is accessible to the server or the application, and the information that is available about a network may be derived from information gathered by a network analyzing product
5 such as a network sniffer, which is a device that is inserted into a segment of the network to monitor the packets that are traveling along the communication line into which the sniffer has been inserted.

The increasing popularity and complexity of Web site and networked database applications that are dependent on the Internet, or on a network
10 using Internet protocols, has further complicated the administrative task, including the task of managing Web site database content and maintaining Web site effectiveness. Company Webmasters and business managers are routinely faced with a wide array of burdensome issues such as the detection, identification, and repair of broken Web sites and links, the monitoring and
15 organization of large volumes of diverse, continuously-changing Web site and networked database content, and the detection and management of Web site problems and congested links. These issues are particularly significant for companies that rely on their Web sites to provide mission-critical information and services to customers and business partners.

-3-

Summary of the Invention

A method and apparatus are provided for evaluating computer resources. To closely simulate the experience of an end user, a software agent executes in a location where the end user would be situated and attempts to exercise a computer resource such as a networked application or a network resource in a simulated transaction in exactly or nearly exactly the same way that the computer resource would be exercised by the end user in a non-simulated transaction. The results of the simulated transaction and of other simulated transactions by other software agents are communicated to a central software system for recordation and analysis.

The method and apparatus may be used with the World-Wide Web or the Internet or with a local area network ("LAN") or a wide area network ("WAN") that extends over a large enterprise and includes a number of LANs connected together in an "intranet."

Different aspects of the invention allow one or more of the following. An end-to-end evaluation can be performed of the ability of an overall system, that includes an application running on a server and a network providing access to the application, to serve an end-user effectively. A pro-active detection system can be provided that detects a problem in the overall system before the problem affects, or is detected by, an end-user. Simulated transactions that are highly representative of non-simulated transactions can be executed using the same computer resources that are used by the non-

-4-

simulated transactions. Existing computer and network hardware already installed in a network can be used highly effectively to evaluate computer resources on the network. Evaluations of computer resources on a network can be performed at effectively any time with little or no disruption to the
5 normal operation of the computer resources or the network.

Other features and advantages will become apparent from the following description, including the drawings, and from the claims.

Brief Description of the Drawings

10 Figs. 1-2 and 7 are block diagrams of computer-based systems.
Figs. 3A-6 are flow diagrams of computer-based procedures.

Detailed Description

A Synthetic Transactions system that simulates the actions of an actual
15 distributed network software application may be used to evaluate computer resource (e.g., application) response time or availability over a computer network. In at least some cases, the evaluations are taken from the perspective of an end user (which may be a human being or a software program), so that the evaluations reflect the application response time and availability that
20 would likely be experienced by the end user.

As described in detail below, a Synthetic Transaction is a simulated application transaction. In at least some cases, it is advantageous if the

-5-

simulated transaction is generated by an intelligent agent software system to evaluate a computer resource. For example, in a typical case concerning client/server technology on a network, a Synthetic Transactions software agent emulates a client side of a client/server transaction so that what happens on the network and on a server is the same, or nearly the same, as what would happen if the transaction were initiated by an actual end user on the client side. The agent may execute at a point on the network that is highly representative of where an actual end user would be situated, so that the path of associated communication across the network closely simulates the path of communication that is, or likely would be, associated with an actual end user.

Fig. 1 illustrates relationships among components of an example implementation ("NextPoint S3") of the Synthetic Transactions system, described in more detail below.

In the Synthetic Transactions system, an application response time is an amount of time between two events in the course of a transaction. Typically, the application response time is an amount of time between a time when a transaction involving an application is started by the software agent and a later time when the transaction is completed (e.g., when the last protocol message has been processed). However, intermediate points in the course of the transaction could be used where appropriate. Typically, the application response time includes any time spent in the transaction due to latencies in the computer network, the application server, any additional backend servers such

-6-

as a database server, as well as time spent by the software agent (e.g., in processing application protocol messages).

An overall application response time measured by a software agent may include multiple components, such as time spent on the network and time
5 spent in the application, which may also include multiple components such as time spent for individual queries in a database application.

One or more of many methods could be used in determining how much of an overall application response time is attributable to network travel time ("network response time") and to time spent in processing by the application
10 ("processing response time"), respectively. In a first example method, an Internet Control Message Protocol Packet Internet Groper ("ICMP Ping") protocol transmission is used to determine the network response time. An ICMP Ping protocol transaction transmission is sent simultaneously, or nearly simultaneously, with an application protocol message to the application server.
15 At the application server, the ICMP Ping protocol transaction transmission is processed at a low level without involving the application software, to issue a response. The amount of time that passes between the sending of the ICMP Ping protocol transmission and the receiving of the response represents the network response time. The processing response time may be determined by
20 subtracting the network response time from the overall application response time.

-7-

In a second example method, a Transmission Control Protocol/Internet Protocol ("TCP/IP") socket connection time is determined and used to determine the network response time. An initial step in executing any TCP/IP network application transaction is the establishment of a TCP/IP socket connection between the client and the server, which establishment is used in the second method to measure the network response time. The establishment of the socket connection does involve the application on the server, because the connection is not successful unless the application accepts the connection, but such involvement does not include hard disk access or back end server components and therefore takes up relatively insignificant time. The network response time is taken to be the TCP/IP socket connection time (i.e., the amount of time needed to establish the socket connection) and the processing response time is determined by subtracting the network response time from the overall application response time.

An advantage of the second example method is that the measurement is made using the same network path that is used in typical non-evaluative transactions between the client and the server. In the case of the first example method, an ICMP Ping transmission may be filtered out by an intermediate network processor such as a firewall, or a network device providing Quality of Service ("QOS") services in place within the computer network may cause the ICMP Ping transmission to be given a priority that is different from the priority that a typical non-evaluative transmission would be given. Such a

-8-

difference in priority may affect the network path followed or the processing time for the ICMP Ping transaction at the server, which would skew the evaluation of the network response time.

5 In a specific implementation, multiple methods, such as both the first and second example methods described above, may be made available for selection and execution.

Application availability is determined based on whether the application responds to a Synthetic Transactions transmission, regardless of the response time. If a response to the Synthetic Transactions transmission is received
10 before a designated expiration time has passed, the application is noted as available. Otherwise, the application is noted as unavailable.

Synthetic Transactions may be executed on a scheduled basis such as a periodic basis, or on user demand. For example, a Synthetic Transaction may repeatedly exercise a specific application function, such as reading a customer
15 record from a database. As a result, a baseline of expectations for the latency of the transaction being simulated may be established, and deviations from the baseline may be interpreted as indicating abnormal functioning of the computer network or the application. In cases in which the network response time and the processing response time are determined separately and
20 corresponding separate baselines are established, the abnormal functioning may be isolated to the network or the application, respectively. See the

-9-

commonly assigned, simultaneously filed United States Patent Application entitled "MANAGING COMPUTER RESOURCES".

Synthetic Transactions may be executed for at least the following network application technologies: ICMP Ping, Dynamic Host Control Protocol ("DHCP"), Domain Name Service ("DNS"), Post Office Protocol 3 ("POP3"), Simple Mail Transfer Protocol ("SMTP"), Standard Query Language ("SQL") Database queries, Hypertext Transport Protocol ("HTTP") web applications, eCommerce, SAP/R3, PeopleSoft, Lotus Notes, Microsoft Exchange and Voice over IP. Synthetic Transactions are not limited to Internet Protocol network applications and may be used for Systems Network Architecture ("SNA"), Internet Protocol Exchange ("IPX"), or Appletalk network applications, for example. All or a part of a network application may be evaluated through the use of Synthetic Transactions. In a typical implementation, a set of Synthetic Transactions is developed to evaluate common and important operations performed by the application. The user may be required to supply parameters that specify a particular instance of the application to be used. For example, the parameters may include the name of a server and a username and password to log into an application on the server.

In at least some cases, it may be advantageous if each set of Synthetic Transactions for an application has characteristics that are particularly appropriate for that application. Although the software agent may provide a common infrastructure for all transactions, the transactions for evaluations of a

-10-

particular application may be tied to that application. For example, SQL scripts are supported by Synthetic Transactions for database queries. The SQL scripts allow for effectively infinite flexibility in the number and type of database related Synthetic Transactions that may be executed. In another example, a Synthetic Transaction to evaluate an ability to look up an Internet Domain Name by the Domain Name Service may have characteristics that render the transaction significantly less useful for another type of evaluation.

In at least some cases, the software agent generates a Synthetic Transaction by creating a set of network protocol messages that are appropriate for an application that is being evaluated. For example, as described in more detail below, to measure response time concerning a Web application, the software agent may build and generate HTTP protocol messages, to emulate a Web browser function. In such a case, a specific Uniform Resource Locator ("URL") identifier for a Web page is encoded in one or more of the HTTP protocol messages and the software agent attempts to retrieve all or a portion of the Web page specified by the URL identifier.

Information generated from a Synthetic Transaction may be temporarily stored by the software agent and may be subsequently sent to a central server for long term storage and for further processing and analysis. The software agent may be controlled by software at the central server, for a purpose such as so that the software agent is specifically directed to generate information needed by the central server for analysis of the network or computer resources.

-11-

For example, software at the central server may determine a set of Synthetic Transactions to be executed and the schedule for the execution of such transactions, and then may apply the set and the schedule to the software agent, which may then execute autonomously to generate the information
5 needed by the software at the central server. (In certain implementations, the software agent may be able to create, specify, or vary the set or the schedule, perhaps in response to varying conditions or inputs.)

Synthetic Transactions may be used to measure throughput, e.g., by keeping track of how much time is required for an amount of data to pass
10 through a particular point in the network.

A specific example of a Synthetic Transaction is described below. Fig. 2 illustrates a network system 50 in which server software ("RT Collector") 52 running on a computer 54 communicates on a network 56 with software agents 58, 60, 62 ("Agents") running on other computers 64, 66, 68. The RT Collector
15 and Agents execute as described below to conduct transactions in connection with evaluating a networked application 70 running on a Web site 72.

The RT Collector is initialized (Figs. 3A-3B) (step 1010), and sets up a timer for the transactions that the RT Collector is to direct the Agents to execute (step 1020). The timer is set to issue a signal at periodic intervals
20 (which may be user defined) when transactions are to be executed.

When the timer signals (step 1030), the RT Collector determines whether the RT Collector is currently in contact with the Agent that is responsible for

-12-

performing a particular transaction (step 1040). If the RT Collector is not in contact with the Agent, the RT Collector opens a TCP socket connection to the Agent on a specified port such as port 9000 (step 1050). (If the RT Collector is unable to establish or re-establish a connection to the Agent, the RT Collector
5 tries to establish the connection every time that the timer signals that the Agent is to perform a transaction.) Once established, the connection between the Agent and the RT Collector is used for all communications between the RT Collector and the Agent until the connection is lost, which may occur due to a network failure or when the client or the server closes the connection.

10 The RT Collector and the Agent determine each other's version numbers as follows (step 1060), so that the RT Collector and the Agent are aware of each other's capabilities. The RT Collector submits, to the Agent, a version request packet that indicates the version number of the RT Collector, and waits a fixed amount of time for the Agent to return another version packet that
15 indicates the version number of the Agent. If the other version packet is not received before the fixed amount of time passes, the RT Collector assumes that the Agent has a version number, e.g., "1.0.0.0.", that corresponds to a basic set of capabilities.

After the version numbers are determined, the RT Collector determines
20 all of the timers that have signaled (step 1070) and identifies a set of transactions corresponding to the timers (step 1080). The RT Collector determines whether a previous instance of any transaction in the set is

-13-

currently being performed by the Agent (step 1090). If so, the RT Collector determines that the previous instance was not performed successfully (which may occur for a reason such as a malfunction in the Agent) (step 1100) and generates an exception indicating that the Agent is unavailable (step 1110).

- 5 Otherwise, the RT Collector submits information for the set of transactions to the Agent (step 1120).

The information for the set of transactions to be performed is sent to the Agent in the form of packets. Each packet has a common header that includes the following information. A transaction ID is included that is to be returned
10 to the RT Collector in response packets from the Agent so that the RT Collector can identify the transaction with which the response is associated. A transaction type ID is included to identify the type of transaction that the Agent is to perform (e.g., Web site transaction, e-commerce transaction, database query transaction). An integer is included to indicate the number of
15 times that the transaction is to be performed.

The remainder of each packet is specific to the transaction type. For example, in the case of a Web site transaction, the remainder may include the URL of a Web page at the Web site and an integer indicating the number of lines to read from the Web page.

- 20 Each response packet sent from the Agent to the RT Collector includes a transaction ID, which in a specific implementation is the first piece of data in the response packet. The remainder of the response packet may be specific to

-14-

the transaction type, and may include a response time value indicating a number of milliseconds of response time. A negative response time value may be used to indicate that an error occurred during performance of the transaction, and may include an error code bit mask that specifies the nature of the error to the RT Collector.

5 The Agent executes as follows (Figs. 4A-4B). The Agent interprets, and executes according to, any customized environment variables that may have been specified, such as variables included on a command line command or in a parameter file maintained on a hosting computer system on which the Agent executes (step 2010).

The Agent waits for a connection attempt on port 9000 (step 2020). When an RT Collector connects, a "Com Thread" thread of execution is established to handle communication between the Agent and the RT Collector (step 2030) and the Agent waits for another connection attempt on port 9000.

15 The Com Thread executes in a loop that includes reading a transaction request packet received from the RT Collector (step 2040), determining the type of transaction that is indicated in the request packet (step 2050), and establishing a Transaction Thread to perform the transaction (step 2060). The Transaction Thread receives, as parameters, a reference to the transaction request packet and a reference to the socket connection, which allows the Transaction Thread to respond directly to the RT Collector independently of other Agent behavior.

-15-

In a specific implementation, the Transaction Thread executes as follows. Transaction-specific information is extracted from the transaction request packet (step 2070), and the transaction is performed to produce results (step 2080). A response packet is produced that includes the results (step 5 2090), and the response packet is sent to the RT Collector (step 2100).

A Web site transaction may be performed as follows (Figs. 5A-5D), in a specific implementation. The URL of a Web site page to be read and a number specifying the number of lines of the page to be read are extracted from the transaction request packet (step 3010). The current time is recorded 10 as a transaction start time (step 3020). An instance of a Java URL class is created based on the URL (step 3030). A network connection is created for the URL (step 3035), and options are set on the network connection, such an option concerning a selectable address or a customization in connection with the computing environment (step 3040). If the user has requested that a 15 password protected HTTP proxy server be used, the encrypted password is specified (step 3050). The URL is designated to be for a Web page that is not to be fetched from any cache (step 3060).

An input stream is opened on the network connection (step 3070), and an input stream reader is opened on the input stream (step 3080). A watchdog 20 timer is created that causes the evaluation to be halted if an expected response to any read request is not received within a specified amount of time, such as 30 seconds (step 3090).

-16-

The following steps 3110, 3120 are performed until the specified number of lines of the Web page are read, an error occurs, or all of the lines of the Web page have been read (step 3100). A line of the Web page is read (step 3110). To prevent the watchdog timer from expiring, the watchdog timer is
5 reset (step 3120).

Once the specified number of lines have been read, the watchdog timer is terminated (step 3130). The current time is recorded as a transaction end time (step 3140). All transaction specific resources are cleared (step 3150).

An elapsed response time is calculated by determining the difference
10 between the transaction end time and the transaction start time (step 3160). A response packet is built that includes the transaction ID that was passed in the request packet and either the elapsed response time or an error code as described above (step 3170). The response packet is sent to the RT Collector (step 3180) and the Transaction Thread is terminated (step 3190).

15 If the watchdog timer expires, the resources associated with the transaction are terminated (step 3200) and the RT Collector is sent a response packet indicating that the transaction timed out (step 3210).

The RT Collector executes as follows upon receiving a response packet from the Agent (Fig. 6). The response packet is appended to a queue of
20 response packets that is configured to hold the response packets of all of the Agents to which the RT Collector submits transaction requests (step 4010). In turn, the response packet is retrieved from the queue of response packets (step

-17-

4020). The transaction with which the response is associated is identified (step 4030) and is determined to be concluded (step 4040). If the results of the transaction indicate an error, an exception is posted (step 4050).

If the hour of the day (or other appropriate interval) has changed, the
5 results of all of the previous hour's instances of the transaction are added to the database (step 4060).

The results of the transaction are stored together with results for other instances of the transaction in the same time period, such as in the current hour (step 4070). If the transaction is executed more often than a particular
10 frequency, such as every 15 minutes, the corresponding results may be averaged (step 4080).

The technique (i.e., the procedures described above) may be implemented in hardware or software, or a combination of both. In at least some cases, it is advantageous if the technique is implemented in computer
15 programs executing on one or more programmable computers, such as a personal computer running or able to run Microsoft Windows 95, 98, or NT, Sun Solaris, LINUX, Unix, or MacIntosh OS, that each include a processor, a storage medium readable by the processor (including volatile and non-volatile memory and/or storage elements), at least one input device such as a
20 keyboard, and at least one output device. Program code is applied to data entered using the input device to perform the method described above and to

-18-

generate output information. The output information is applied to one or more output devices such as a display screen of the computer.

In at least some cases, it is advantageous if each program is implemented in a high level procedural or object-oriented programming language such as C++ or Java to communicate with a computer system. However, the programs can be implemented in assembly or machine language, if desired. In any case, the language may be a compiled or interpreted language.

In at least some cases, it is advantageous if each such computer program is stored on a storage medium or device, such as ROM or magnetic diskette, that is readable by a general or special purpose programmable computer for configuring and operating the computer when the storage medium or device is read by the computer to perform the procedures described in this document. The system may also be considered to be implemented as a computer-readable storage medium, configured with a computer program, where the storage medium so configured causes a computer to operate in a specific and predefined manner.

Other embodiments are within the scope of the following claims. For example, transactions may also be performed for e-commerce applications or database queries or for evaluations of network throughput or mass storage performance. In the case of database queries, a database transaction system

-19-

can be arranged in which SQL statements are submitted by a software agent to a remote database server on a periodic basis such as every five minutes.

The software agent may execute tasks such as breaking down the test results into components including the connection time, and automatically
5 executing a transaction multiple times at predetermined intervals in response to a single directive from the RT Collector. Execution threads may be arranged so that multiple RT Collectors may be served by a single software agent, and a single RT Collector may be served by multiple software agents. The software agent may acknowledge receipt of a request packet from the RT
10 Collector by sending an acknowledgement packet to the RT Collector. The software agent may have a class structure, such as the structure illustrated in Fig. 7, within which processing tasks are delegated. The class structure may include a facility for processing the results of a transaction, and perhaps also a facility that generates an alarm, an alert, or an exception when the processing
15 indicates that the results merit drawing attention to a computer resource involved in the transaction.

Transactions may be grouped by transaction IDs. Java based classes of distributed processes may be used, which may provide significant platform independence.

20 The software agent may maintain a data storage cache in which to store the results of multiple transactions, which may reduce the amount or frequency of communication between the software agent and the RT Collector

-20-

and may allow the software agent to operate autonomously during a period in which the RT Collector cannot communicate with the software agent. If the cache becomes full or nearly full, information in the cache may be discarded on a first in, first out basis.

- 5 The RT Collector may serve as a socket client. Java archive objects, SNMP Management Information Base ("MIB"), Extensible Markup Language ("XML") page description technology, or Push technology may be used, and may allow the software agent to supply results to the RT Collector or to another destination on a Push technology basis.
- 10 The transaction specific information supplied by the RT Collector may be derived from information that the RT Collector extracts from a database of evaluation transaction setup information.

-21-

What Is Claimed Is:

1. A method comprising:

causing a transaction to be executed that emulates an end-user's
interaction with a software application, wherein the transaction includes
5 communication with the software application across a network.

2. The method of claim 1, further comprising:

causing a message to be transmitted, to a point where the software
application is executing, from an end-user point on the network.

10

3. The method of claim 1, further comprising:

causing a message to be transmitted, from a point where the software
application is executing, to an end-user point on the network.

15

4. The method of claim 1, wherein the software application includes a
Web server application.

5. The method of claim 4, wherein:

causing the transaction to be executed includes emulating a function of

20 a Web browser.

-22-

6. The method of claim 1, wherein the network includes client/server technology and the transaction includes a client/server transaction.

7. The method of claim 6, further comprising:

5 emulating the client side of the client/server transaction.

8. The method of claim 1, further comprising

attempting to retrieve specified information from the software application.

10

9. The method of claim 1, further comprising

attempting to retrieve information from a Web page supplied by the software application.

15

10. The method of claim 1, further comprising:

supplying login information to the software application.

11. The method of claim 1, wherein the software application includes a database application.

20

12. The method of claim 11, further comprising:

attempting to retrieve information from the database application.

-23-

13. The method of claim 1, further comprising:
determining a response time in the transaction.

14. The method of claim 13, further comprising:
5 determining an application processing time component of the response
time.

15. The method of claim 13, further comprising:
determining a network processing time component of the response time.

10 16. The method of claim 1, further comprising:
attempting to establish a TCP/IP connection; and
based on a result of the attempt, determining a network processing
time.

15 17. The method of claim 1, further comprising:
transmitting a ping message; and
based on a result of the transmission, determining a network processing
time.

20 18. The method of claim 1, wherein the emulating includes emulating a
TCP/IP client.

-24-

19. The method of claim 1, further comprising:
based on the transaction, determining the availability of the software application.

5 20. The method of claim 19, further comprising:
if a response is not received timely from the software application,
determining that the software application is unavailable.

10 21. The method of claim 1, further comprising:
causing multiple instances of the transaction to be executed according to
a schedule.

15 22. The method of claim 21, further comprising:
deriving a performance baseline from results of the multiple instances.

23. The method of claim 1, further comprising:
causing, from a central location, multiple emulation transactions to be
executed.

20 24. The method of claim 23, wherein at least two of the transactions
originate from two different respective locations on the network.

-25-

25. The method of claim 24, further comprising:

storing information derived from a result of at least one of the transactions at the respective originating location for later transmission to the central location.

5

26. The method of claim 1, further comprising:

transmitting a data packet to the software application, the data packet including information that identifies the transaction.

10

27. A method comprising:

generating a simulated computer communication at a point in a computer network;

using a non-simulated network resource to deliver the simulated computer communication to a computer resource; and

15

evaluating the computer resource based on a result of the delivery of the simulated computer communication to the computer resource.

28. The method of claim 27, wherein the point in the computer network includes an end-user point.

20

29. The method of claim 27, further comprising:

-26-

issuing, from a central source, a directive to generate the simulated computer communication; and

receiving, at the central source, a report on the result of the delivery of the simulated computer communication to the computer resource.

5

30. Computer software, residing on a computer-readable storage medium, comprising a set of instructions for use in a computer system to cause the computer system to:

10 cause a transaction to be executed that emulates an end-user's interaction with a software application, wherein the transaction includes communication with the software application across a network.

31. Computer software, residing on a computer-readable storage medium, comprising a set of instructions for use in a computer system to
15 cause the computer system to:

generate a simulated computer communication at a point in a computer network;

use a non-simulated network resource to deliver the simulated computer communication to a computer resource; and

20 evaluate the computer resource based on a result of the delivery of the simulated computer communication to the computer resource.

-27-

32. A system comprising:

a system that causes a transaction to be executed that emulates an end-user's interaction with a software application, wherein the transaction includes communication with the software application across a network.

5

33. A system comprising:

a generator that generates a simulated computer communication at a point in a computer network;

a utilizer that uses a non-simulated network resource to deliver the
10 simulated computer communication to a computer resource; and
an evaluator that evaluates the computer resource based on a result of the delivery of the simulated computer communication to the computer resource.

1/11

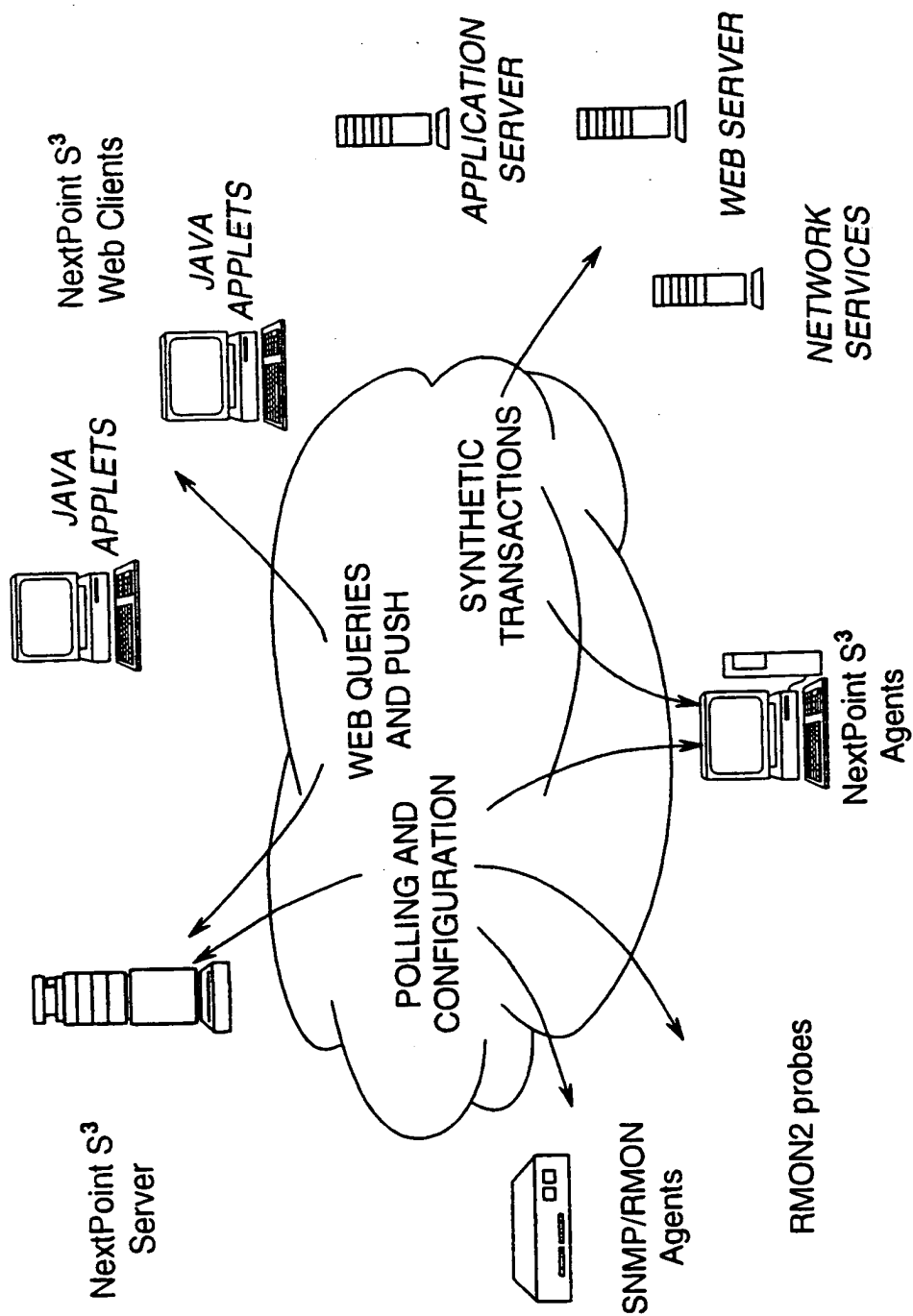


FIG. 1

2/11

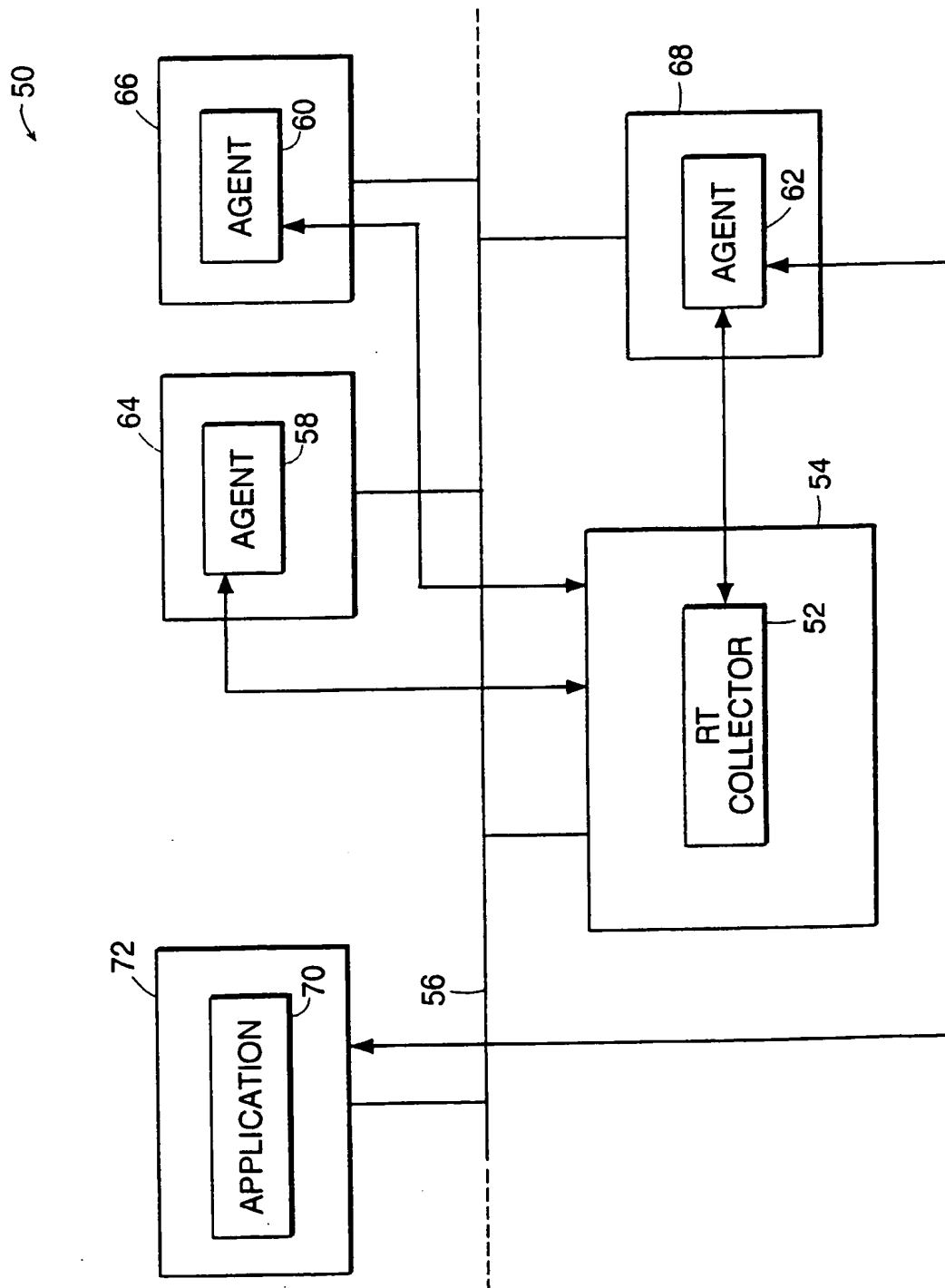


FIG. 2

3/11

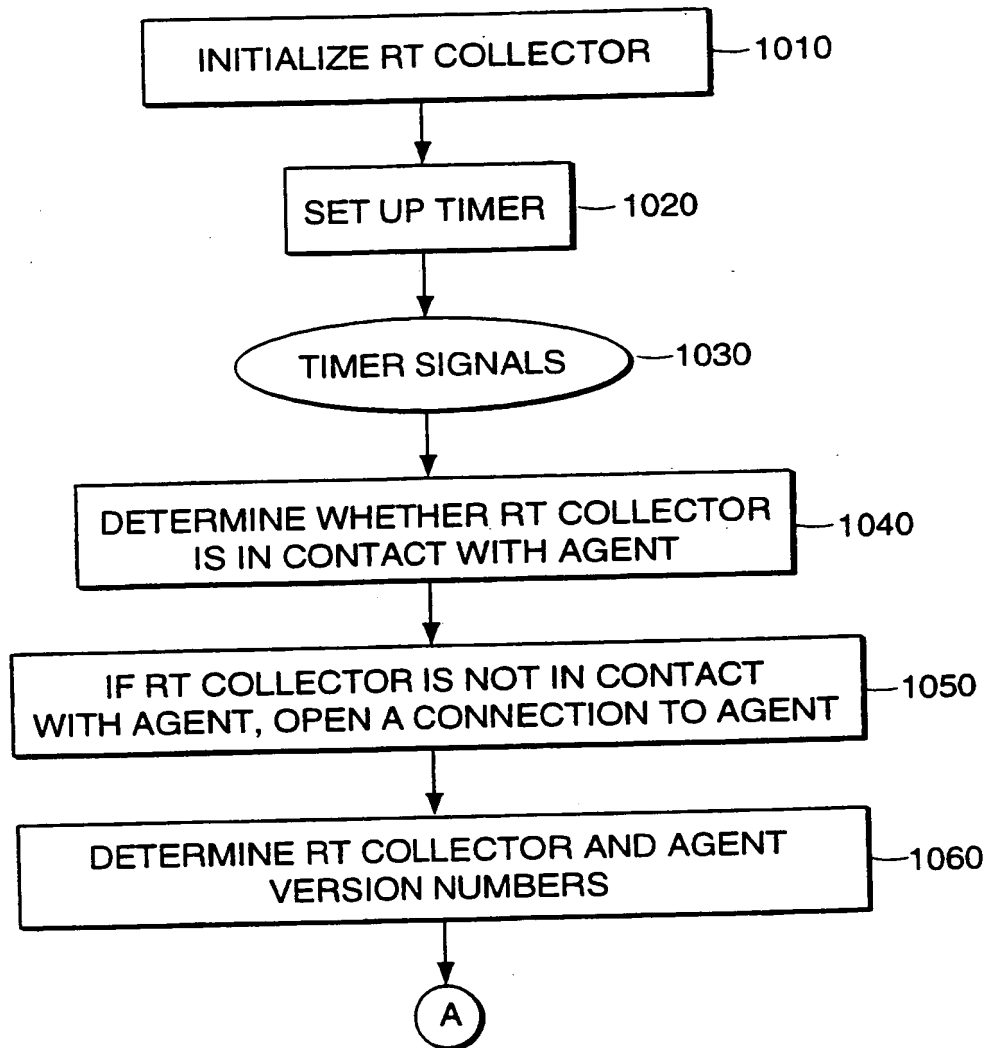


FIG. 3A

4/11

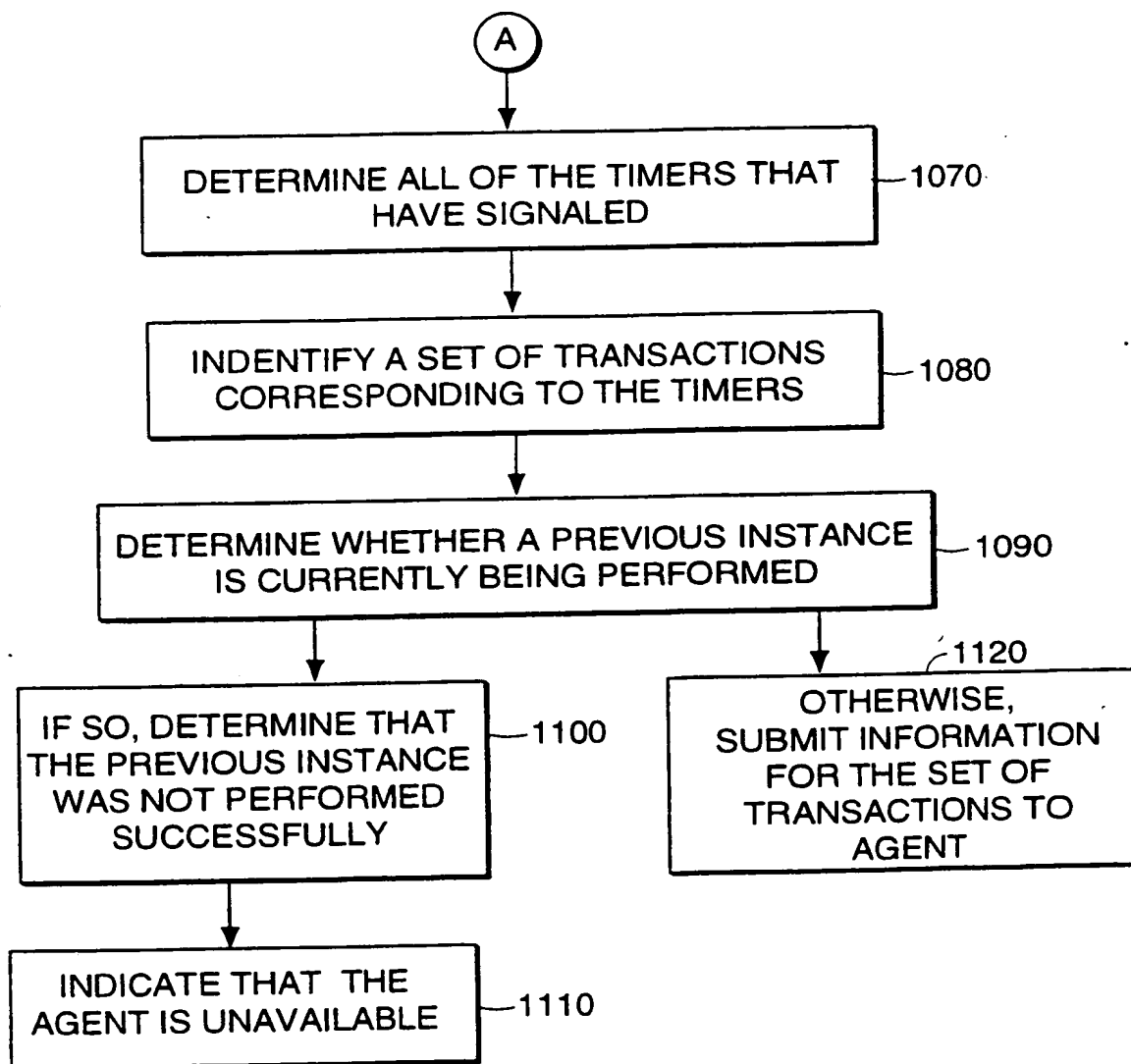


FIG. 3B

5/11

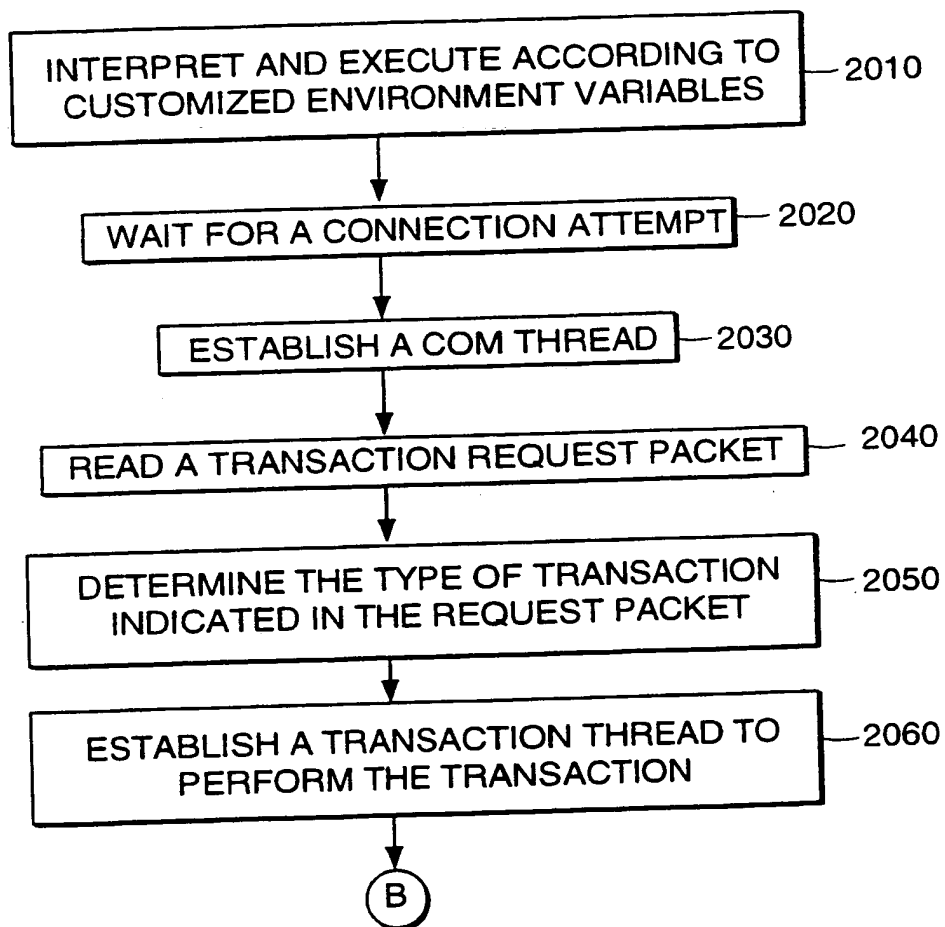


FIG. 4A

6/11

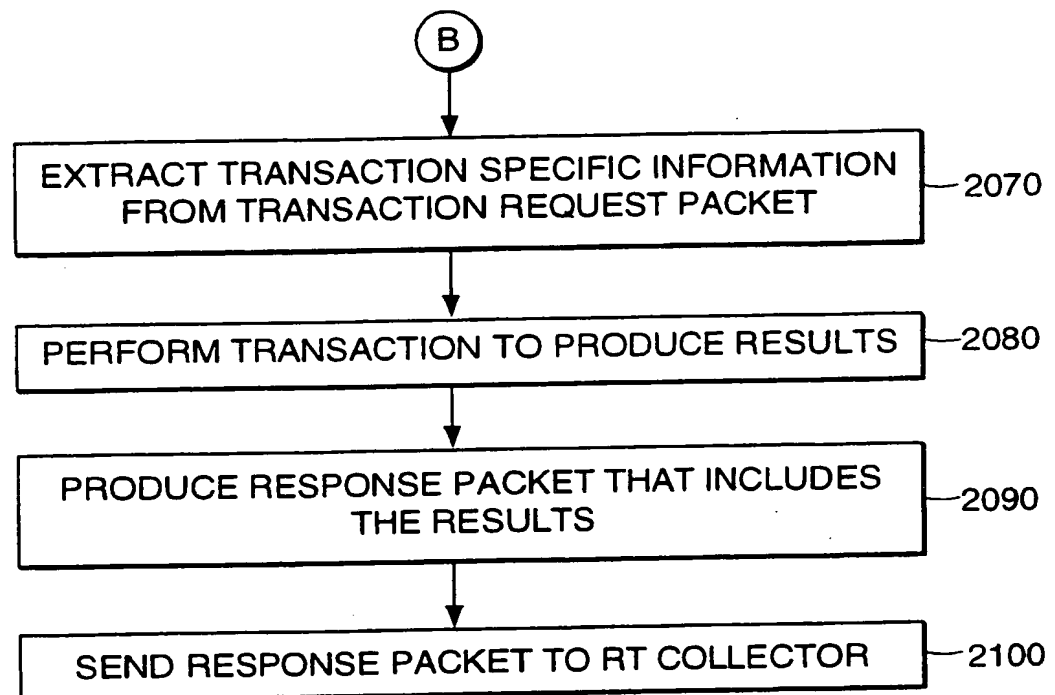


FIG. 4B

7/11

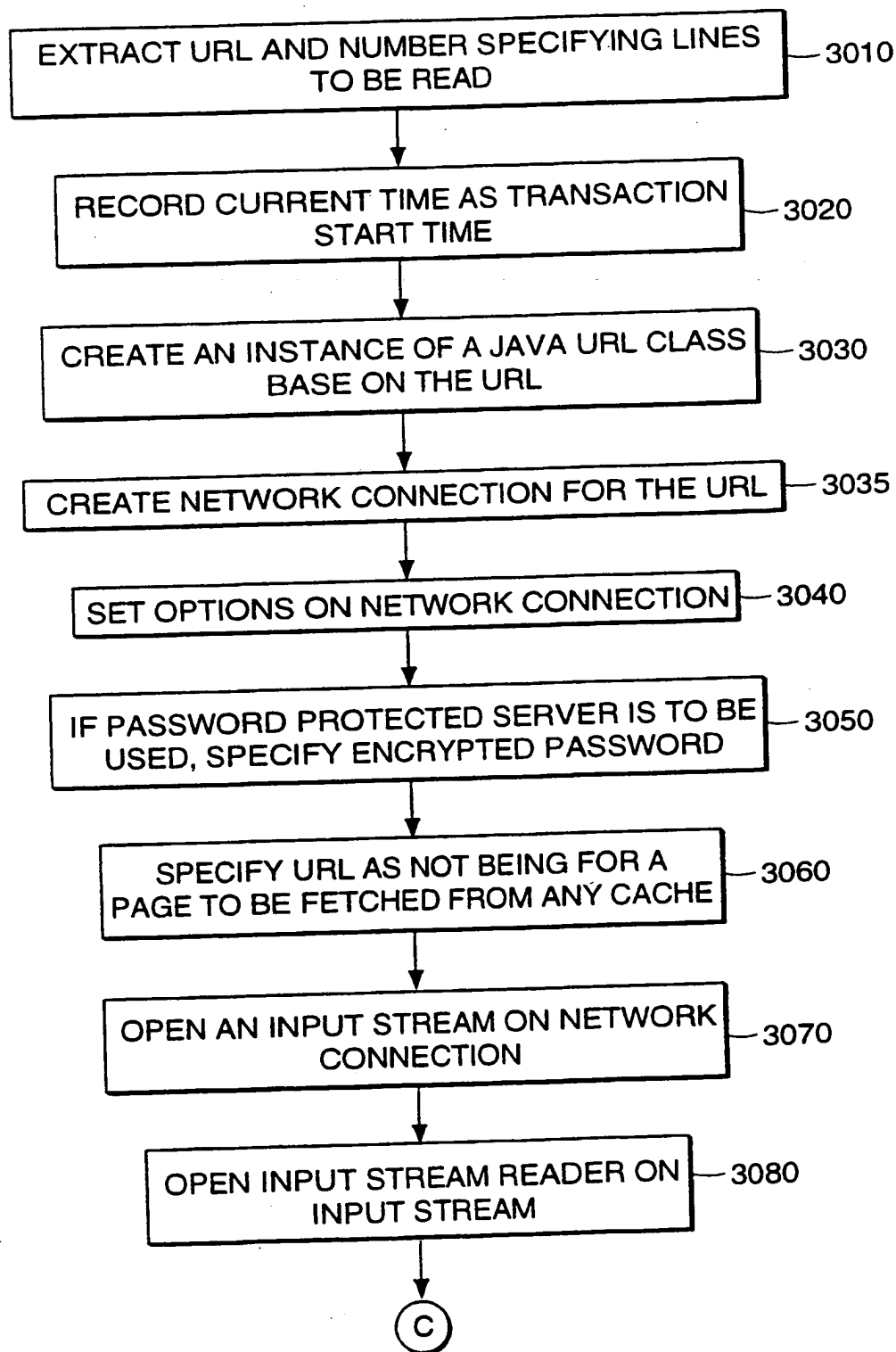


FIG. 5A

8/11

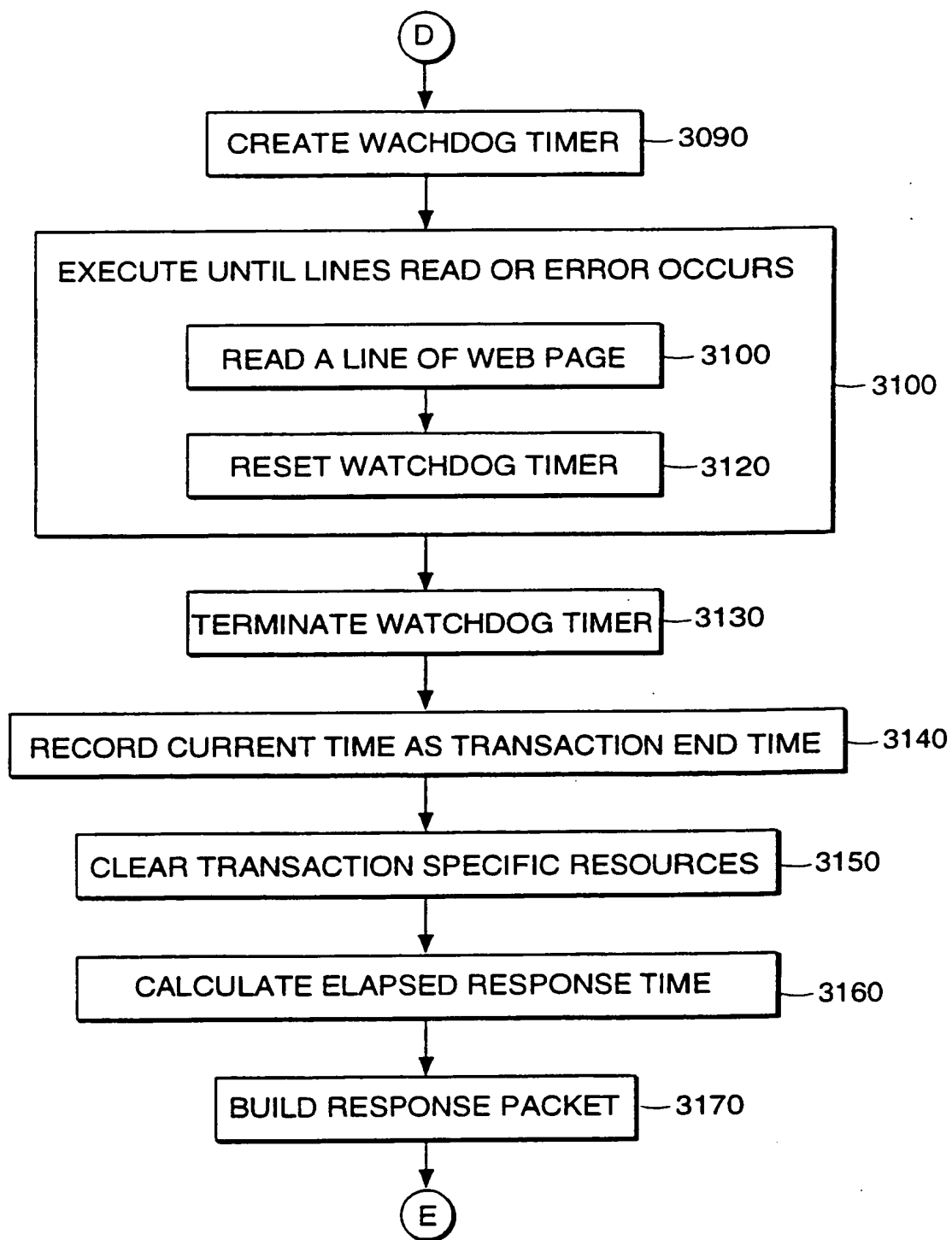


FIG. 5B

9/11

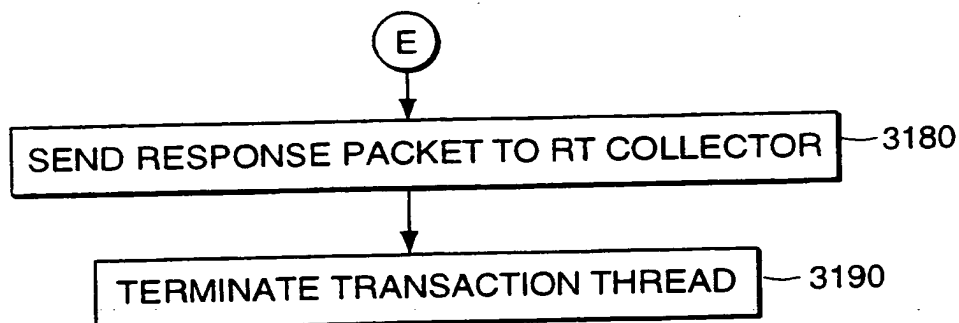


FIG. 5C

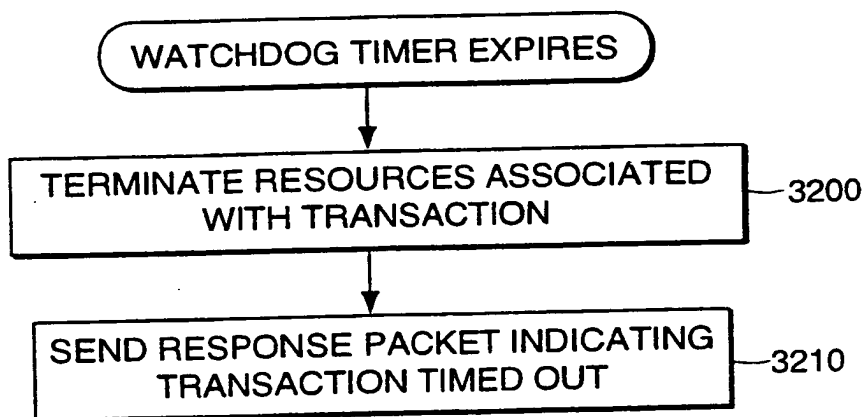


FIG. 5D

10/11

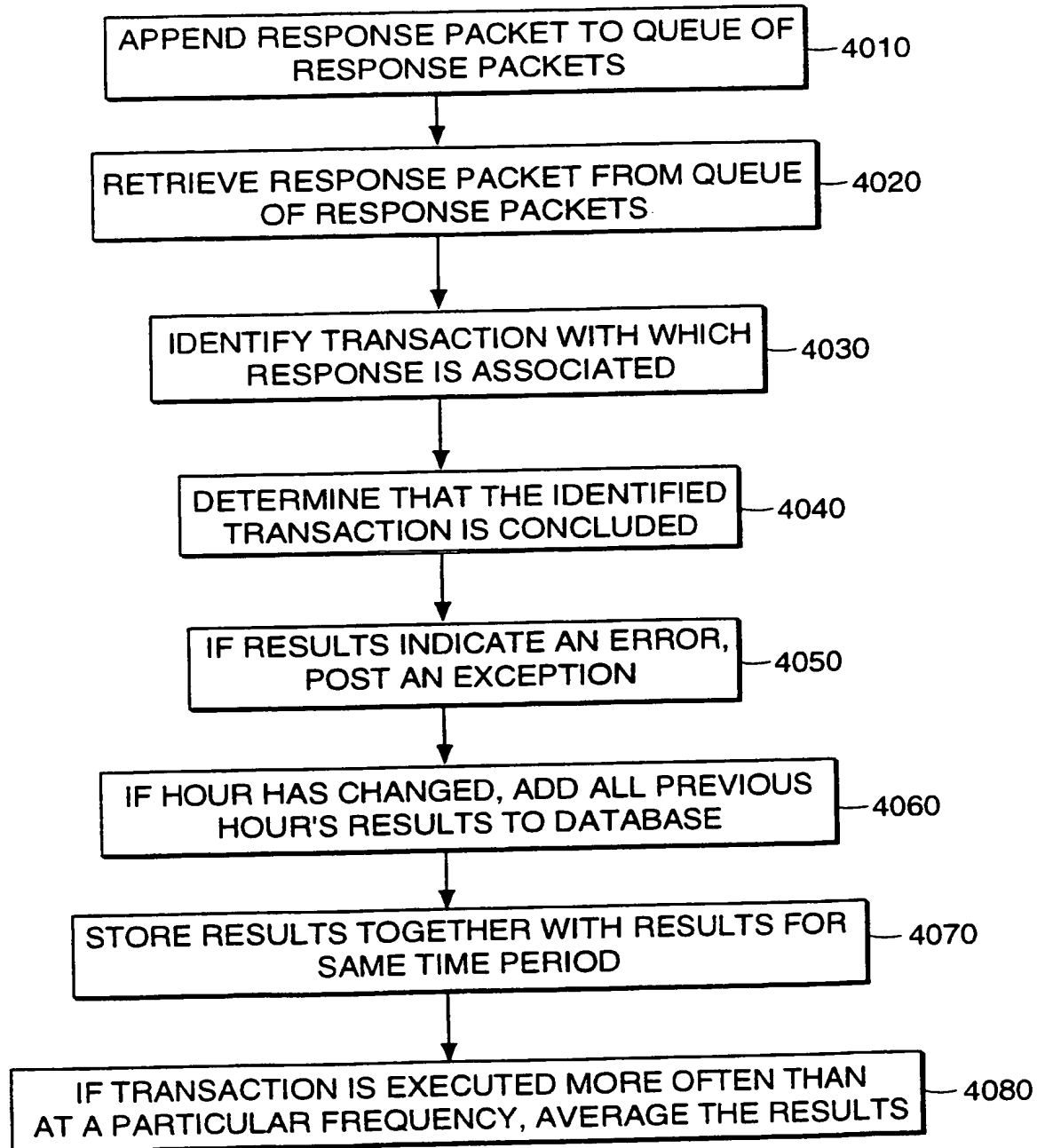


FIG. 6

11/11

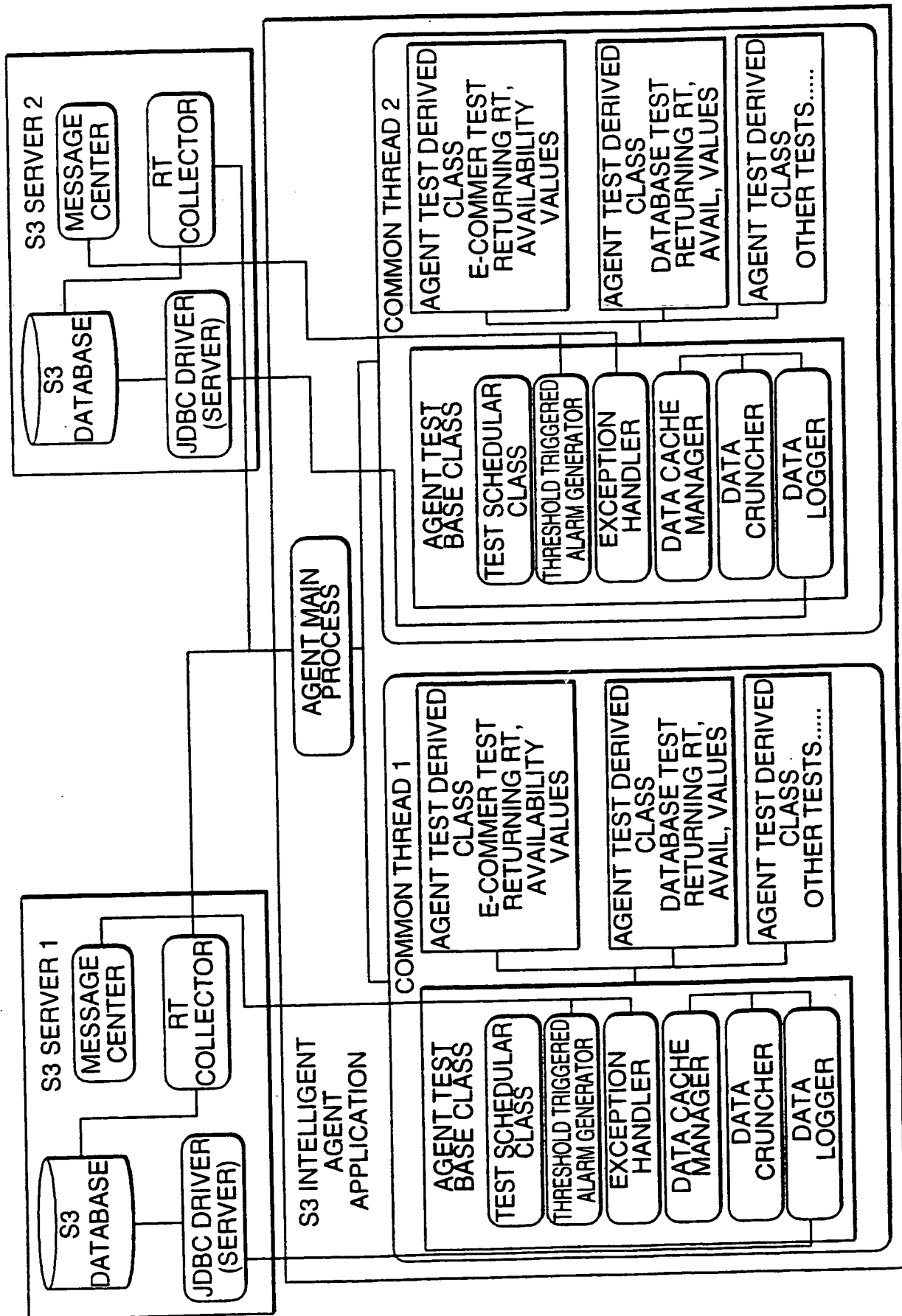


FIG. 7

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 99/22787

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 G06F11/34 H04L12/26

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 G06F H04L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5 812 780 A (CHEN JOHN YUN-KUANG ET AL) 22 September 1998 (1998-09-22) column 1, line 1 -column 9, line 14 column 13, line 45 -column 14, line 38	1-4, 6-8, 10-13, 18-24, 26-33
X	US 5 577 197 A (BECK LONNIE P) 19 November 1996 (1996-11-19) the whole document	1-3, 6-8, 11-15, 19-21, 23-33
X	GB 2 307 318 A (MITSUBISHI ELECTRIC CORP) 21 May 1997 (1997-05-21) claim 1	1-3, 6-8, 11-13, 21-33

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"A" document member of the same patent family

Date of the actual completion of the international search

21 February 2000

Date of mailing of the international search report

06/03/2000

Name and mailing address of the ISA
European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3018

Authorized officer

Herreman, G

INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 99/22787

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	PATENT ABSTRACTS OF JAPAN vol. 006, no. 137 (P-130), 24 July 1982 (1982-07-24) & JP 57 059257 A (AGENCY OF IND SCIENCE & TECHNOL), 9 April 1982 (1982-04-09) abstract	1-3,6,7, 13-15, 21,23, 24,27-33
A	BORELLA M S ET AL: "WANDS: Wide-Area Network Delay Simulator" PROCEEDING OF THE IEEE 7TH INTERNATIONAL WORKSHOP ON NETWORK AND OPERATING SYSTEMS SUPPORT FOR DIGITAL AUDIO AND VIDEO (CAT. NO.97TH8259), PROCEEDINGS OF 7TH INTERNATIONAL WORKSHOP ON NETWORK AND OPERATING SYSTEM SUPPORT FOR DIGITAL AUDIO AND VIDEO, pages 85-92, XP002131053 1997, New York, NY, USA, IEEE, USA ISBN: 0-7803-3799-9 page 87, paragraph 3.1CRAWLER	1,4,5,9
A	LARSEN A K: "ALL EYES ON IP TRAFFIC. NEW APPS CAN MONITOR INTERNET AND INTRANET TRAFFIC, BUT DO THEY DELIVER ENOUGH DATA TO HOLD ISPS TO THEIR PROMISES?" DATA COMMUNICATIONS,US,MCGRAW HILL. NEW YORK, vol. 26, no. 4, 21 March 1997 (1997-03-21), pages 54,56-60,62, XP000659549 ISSN: 0363-6399 page 59, middle column, line 13 -page 60, left-hand column, line 14	16,17

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 99/22787

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
US 5812780	A	22-09-1998	NONE	
US 5577197	A	19-11-1996	NONE	
GB 2307318	A	21-05-1997	JP 9138777 A US 5889955 A	27-05-1997 30-03-1999
JP 57059257	A	09-04-1982	JP 1323657 C JP 60053891 B	27-06-1986 27-11-1985

This Page Blank (uspto)